

Analysis of Asymptotic Preserving DG-IMEX Schemes for Linear Kinetic Transport Equations in a Diffusive Scaling

Juhi Jang¹ Fengyan Li² Jing-Mei Qiu³ Tao Xiong⁴

Abstract

In this paper, some theoretical aspects will be addressed for the asymptotic preserving DG-IMEX schemes recently proposed in [10] for kinetic transport equations under a diffusive scaling. We will focus on the methods that are based on discontinuous Galerkin (DG) spatial discretizations with the P^k polynomial space and a first order IMEX temporal discretization, and apply them to two linear models: the telegraph equation and the one-group transport equation in slab geometry. In particular, we will establish uniform numerical stability with respect to Knudsen number ε using energy methods, as well as error estimates for any given ε . When $\varepsilon \rightarrow 0$, a rigorous asymptotic analysis of the schemes is also obtained. Though the methods and the analysis are presented for one dimension in space, they can be generalized to higher dimensions directly.

Keywords: Kinetic transport equations; Asymptotic preserving; High order discontinuous Galerkin method; IMEX; Stability analysis; Error estimate

1 Introduction

Kinetic theory is at the center of multi-scale modeling connecting the invisible microscopic models with the macroscopic hydrodynamic models. In particular, when the mean free path of particles is sufficiently small, the system is close to the equilibrium state and a macroscopic model is a good approximation to the kinetic equation. Building a passage from kinetic to macroscopic models is a very interesting problem and there has been a lot of mathematical progress over the decades [1, 26].

Designing accurate and efficient numerical schemes for kinetic equations with a broad range of Knudsen number ε has been an active research area for more than two decades. Among many multi-scale approaches, asymptotic preserving (AP) methods are known to be able to effectively deal with multi-scales and capture hydrodynamic macro-scale limits. Specifically, the schemes are designed to mimic the asymptotic limit from the kinetic to the hydrodynamic models on the PDE level. As a result, the scheme in the limit of $\varepsilon \rightarrow 0$ becomes a consistent discretization of the limiting macro-scale equations.

AP schemes have been intensively studied in different settings and under different scalings, for example for stationary problems [18, 17, 9] and for time dependent problems with hyperbolic and diffusive scalings [11, 24]. Below, to our best knowledge, we briefly review some existing AP methods for kinetic equations in the diffusive limit. It was firstly shown in [12, 23] that an improper treatment of spatial discretizations even with a stable implicit time discretization may fail to capture the correct asymptotic limit with an under-resolved computational mesh. In [13, 23, 14], some proper splitting between the convection and stiff source terms was introduced for AP properties. Later, AP schemes were designed via different multi-scale approaches, e.g. by a standard perturbation procedure [15], by moment closure approaches [5], by a micro-macro decomposition of kinetic transport equations [19] and by projective integration [16]. AP schemes have also been designed with different high order discretization strategies such as the discontinuous Galerkin framework [22], weighted essentially non-oscillatory (WENO) methods as well as globally stiffly accurate implicit-explicit (IMEX) schemes [3, 4]. Despite the fact that much computational effort was made in designing various AP schemes, rigorous proofs for uniform stability, error estimates and AP properties for fully discrete schemes are relatively rare. For AP schemes based on the micro-macro decomposition, there are some theoretical results. In [19], a von Neumann analysis was conducted for numerical stability of a first order AP scheme applied to the two-velocity telegraph equation. Stability and error estimate for more general problems were obtained in [20] based on energy methods. In the setting of stationary problems, a rigorous asymptotic analysis was presented for an upwind

¹Department of Mathematics, University of California Riverside, Riverside, CA 92521. E-mail: juhijang@math.ucr.edu. Supported in part by NSF grants DMS-0908007 and DMS-1212142.

²Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180. E-mail: lif@rpi.edu. Supported in part by NSF CAREER award DMS-0847241 and NSF DMS-1318409.

³Department of Mathematics, University of Houston, Houston, 77204. E-mail: jingqiu@math.uh.edu. The third and fourth authors are supported in part by Air Force Office of Scientific Computing YIP grant FA9550-12-0318, NSF grant DMS-1217008 and University of Houston.

⁴Department of Mathematics, University of Houston, Houston, 77204. E-mail: txiong@math.uh.edu

discontinuous Galerkin discretization, together with the convergent property of the limiting schemes under mesh refinement, when solving the radiative transport equation [9].

Recently, a family of high order schemes was proposed in [10] for several linear and nonlinear *discrete-velocity* kinetic transport equations under a diffusive scaling. The schemes are defined for a reformulated set of equations which is obtained from a micro-macro decomposition of the problem just as in [19], with the idea originally proposed in [21]. Based on such reformulation, discontinuous Galerkin (DG) spatial discretizations of arbitrary order of accuracy are applied with suitable numerical fluxes, and in time, we employ globally stiffly accurate high order IMEX Runge-Kutta (RK) methods [3] equipped with a carefully chosen implicit-explicit strategy. A formal asymptotic analysis shows that the proposed methods, as Knudsen number ε goes to 0, become consistent high order discretizations for the limiting macro-scale equations. Numerical results presented in [10] also demonstrate the stability and high order accuracy of the proposed schemes when ε is of order 1 and in the limit of ε going to 0.

The current paper follows up the work in [10] and addresses some theoretical aspects of the proposed methods. In particular, we will establish uniform stability, error estimates, and perform a rigorous asymptotic analysis for the fully discrete scheme when DG spatial discretizations using the P^k polynomial space are coupled with a first order IMEX time discretization. Two families of linear kinetic transport equations are considered: the two-velocity telegraph equation, and the one-group transport equation in slab geometry with a continuous velocity field, for which the method in [10] can be directly formulated and applied. In this work, uniform time step constraint with respect to ε is established for numerical stability by using energy methods. More specifically, for the DG method with the P^0 polynomial space, optimal time step restriction is achieved as in [19, 20], namely $\Delta t = O(h^2)$ in the diffusive regime with $\varepsilon \ll 1$, and $\Delta t = O(\varepsilon h)$ in the convective regime with $\varepsilon = O(1)$; for the method with the P^k ($k \geq 1$) polynomial space, the time step restriction is $\Delta t = O(h^2)$, which is not the most desired condition in the convective regime. It is expected that extending the stability analysis to high order IMEX schemes can relieve the time step constraint for the method with the P^k ($k \geq 1$) space [29]. When higher order temporal discretizations are used, most techniques to analyze the spatial discretizations can be carried over, yet new difficulties will arise related to the high order time integrations [28, 29]. Such analysis is a subject of our future investigation. Based on numerical stability and approximation properties of the discrete spaces, error estimates are further established for the schemes with different choices of numerical fluxes. The results confirm the high order accuracy in space and the first order accuracy in time of the methods. By using the weakly sequential compactness of Hilbert space (which in the present work is either the finite-dimensional discrete space or $L^2(\Omega_v)$ with Ω_v given in Section 2), we also prove that the proposed schemes will converge to some consistent discretizations of the limiting heat equation when $\varepsilon \rightarrow 0$. Though the methods and the analysis are presented here and in [10] for one dimension in space, they can be generalized to higher dimensions directly.

The paper is organized as follows. In Section 2, we introduce the kinetic transport equation in a diffusive scaling, provide its micro-macro decomposition and diffusive limit, and review a class of DG-IMEX schemes introduced in [10]. In Section 3, stability analysis, error estimates, and a rigorous asymptotic analysis are carried out for the schemes and discussed in various settings.

2 Formulation

We consider the following linear kinetic transport equation in a diffusive scaling

$$\varepsilon f_t + v \partial_x f = \frac{1}{\varepsilon} (\langle f \rangle - f) \quad (2.1)$$

with the initial data f_0 and spatially periodic boundary conditions, where $f = f(x, v, t)$ is the distribution function of particles that depends on time $t > 0$, position $x \in \Omega_x \subset \mathbb{R}$, and velocity $v \in \Omega_v$. The parameter $\varepsilon > 0$ measures the distance of the system to the equilibrium state and it can be regarded as the mean free path of the particles; when ε is small, the system is close to equilibrium; when ε is large, the system is far from equilibrium. The operator $\langle f \rangle - f$ is the normalized scattering operator. Here $\langle f \rangle = \int_{\Omega_v} f d\mu$, and μ is a measure associated with the velocity space Ω_v and it will be specified next for each model. In this paper, we discuss two important families of the problem (2.1): the telegraph equation and the one-group transport equation in slab geometry.

- Telegraph equation is a discrete-velocity kinetic model with $\Omega_v = \{-1, 1\}$, namely, v either takes value -1

or 1, and μ is a discrete measure on $\{-1, 1\}$ such that

$$\langle f \rangle = \int_{\Omega_v} f d\mu := \frac{1}{2}(f(x, v = -1, t) + f(x, v = 1, t)).$$

• The one-group transport equation in slab geometry is a kinetic equation (2.1) where the velocity space, $\Omega_v = [-1, 1]$, is continuous. In addition, $d\mu = \frac{1}{2}dv$, with dv being the standard Lebeque measure on Ω_v , and

$$\langle f \rangle = \int_{\Omega_v} f d\mu := \frac{1}{2} \int_{-1}^1 f(v) dv.$$

The scattering operator on the right side of (2.1) can be of more general form: $(Lf)(v) = \int_{-1}^1 s(v, v')(f(v') - f(v)) dv'$ where the kernel s satisfies $0 < s_m \leq s(v, v') \leq s_M$ for all $v, v' \in [-1, 1]$, $\int_{-1}^1 s(v, v') dv' = 1$, and $s(v, v') = s(v', v)$. For such kernel s , one can deduce that $\langle L\phi \rangle = 0$ for all $\phi \in L^2[-1, 1]$ and $\langle \phi L\phi \rangle \leq -2s_m \langle \phi^2 \rangle$ for $\phi \in \mathcal{N}(L)^\perp$ [2]. Of course, the simplest case of such s is $s(v, v') = 1/2$, which is the case of (2.1). We refer to [6] for more detailed discussions on the linear transport equation and [20] for the one-group transport equation in slab geometry.

It turns out that dealing with discrete velocity or continuous velocity does not affect much the formulation of the numerical methods and the theoretical results. We will treat both cases in a uniform setting and point out the differences when necessary. Let us consider the Hilbert space $L^2(\Omega_v; d\mu)$ in v variable with the inner product: $\langle f, g \rangle := \int f g d\mu = \langle fg \rangle$, and let Π be the orthogonal projection operator onto $\text{Span}(1)$, defined as $\Pi : f \mapsto \Pi f = \langle f \rangle$. Let $\rho := \Pi f = \langle f \rangle$ denote the macroscopic density for f and we write

$$f = \langle f \rangle + \varepsilon g = \rho + \varepsilon g \tag{2.2}$$

where $\langle g \rangle = 0$. We recall the micro-macro formulation for (2.1) in [10], motivated by [19, 21],

$$\begin{aligned} \partial_t \rho + \partial_x \langle vg \rangle &= 0, \\ \partial_t g + \frac{1}{\varepsilon} (\mathbf{I} - \Pi)(v \partial_x g) + \frac{1}{\varepsilon^2} v \partial_x \rho &= -\frac{1}{\varepsilon^2} g. \end{aligned} \tag{2.3}$$

The operator \mathbf{I} used here is the identity operator. It is easy to verify that $\partial_t \rho = \partial_x (\langle v^2 \rangle \partial_x \rho) + O(\varepsilon)$ which leads to the linear diffusion equation as $\varepsilon \rightarrow 0$. We note that $\langle v^2 \rangle = 1$ for the telegraph equation and $\langle v^2 \rangle = 1/3$ for the one-group transport equation in slab geometry.

Recently in [10], a family of asymptotic preserving methods were proposed for the telegraph equation based on its micro-macro decomposition (2.3). The methods are of formal high order accuracy in both space and time. They involve discontinuous Galerkin (DG) spatial discretizations and globally stiffly accurate implicit-explicit (IMEX) Runge-Kutta methods in time. In the limit of $\varepsilon \rightarrow 0$, a formal asymptotic analysis [10] shows that the limiting schemes are consistent high order discretizations for the limiting linear heat equation. Though not discussed in [10], both the methods and the formal analysis can be naturally extended to the one-group transport equation in slab geometry where the velocity field is continuous. Next we will present the formulation of the methods applied to (2.1), before establishing numerical stability, error estimates, and a rigorous asymptotic analysis in the following section.

Let's first introduce some notation. Start with $\{x_{i+\frac{1}{2}}\}_{i=0}^{i=N}$, a partition of $\Omega_x = [x_{\min}, x_{\max}]$. Here $x_{\frac{1}{2}} = x_{\min}$, $x_{N+\frac{1}{2}} = x_{\max}$, each element is denoted as $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ with its length h_i , and $h = \max_i h_i$. Given any non-negative integer k , we define a finite dimensional discrete space

$$U_h^k = \{u \in L^2(\Omega_x) : u|_{I_i} \in P^k(I_i), \forall i\}, \tag{2.4}$$

where the local space $P^k(I)$ consists of polynomials of degree at most k on I . Note functions in U_h^k are piecewise-defined, we further denote jump and average of u at $x_{i+\frac{1}{2}}$, $\forall i$, as $[u]_{i+\frac{1}{2}} = u(x_{i+\frac{1}{2}}^+) - u(x_{i+\frac{1}{2}}^-)$ and $\{u\}_{i+\frac{1}{2}} = \frac{1}{2}(u(x_{i+\frac{1}{2}}^+) + u(x_{i+\frac{1}{2}}^-))$, respectively. Here $u(x^\pm) = \lim_{\Delta x \rightarrow 0^\pm} u(x + \Delta x)$, and we also use $u_{i+\frac{1}{2}} = u(x_{i+\frac{1}{2}})$, $u_{i+\frac{1}{2}}^\pm = u(x_{i+\frac{1}{2}}^\pm)$, $\forall i$.

With the same DG spatial discretization proposed in [10], a family of semi-discrete methods are given below for the micro-macro system (2.3). Look for $\rho_h(\cdot, t), g_h(\cdot, v, t) \in U_h^k$, such that $\forall \phi, \psi \in U_h^k$,

$$(\partial_t \rho_h, \phi) + a_h(g_h, \phi) = 0, \quad (2.5a)$$

$$(\partial_t g_h, \psi) + \frac{1}{\varepsilon} b_{h,v}(g_h, \psi) - \frac{v}{\varepsilon^2} d_h(\rho_h, \psi) = -\frac{1}{\varepsilon^2} (g_h, \psi), \quad (2.5b)$$

where

$$a_h(g_h, \phi) = - \sum_i \int_{I_i} \langle v g_h \rangle \partial_x \phi dx - \sum_i \langle \widehat{v g_h} \rangle_{i-\frac{1}{2}} [\phi]_{i-\frac{1}{2}}, \quad (2.6a)$$

$$b_{h,v}(g_h, \psi) = ((\mathbf{I} - \Pi) \mathcal{D}_h(g_h; v), \psi) = (\mathcal{D}_h(g_h; v) - \langle \mathcal{D}_h(g_h; v) \rangle, \psi), \quad (2.6b)$$

$$d_h(\rho_h, \psi) = \sum_i \int_{I_i} \rho_h \partial_x \psi dx + \sum_i \widehat{\rho}_{h,i-\frac{1}{2}} [\psi]_{i-\frac{1}{2}}. \quad (2.6c)$$

Here and below, the standard inner product (\cdot, \cdot) for the $L^2(\Omega_x)$ space is used, see e.g. the first term in (2.5a) and in (2.5b). The function $\mathcal{D}_h(g_h; v)$ in (2.6b) belongs to U_h^k , and it is defined based on an upwind discretization of $v \partial_x g$ within the DG framework,

$$(\mathcal{D}_h(g_h; v), \psi) := - \sum_i \left(\int_{I_i} v g_h \partial_x \psi dx \right) - \sum_i \langle \widehat{v g_h} \rangle_{i-\frac{1}{2}} [\psi]_{i-\frac{1}{2}}, \quad \psi \in U_h^k, \quad (2.7)$$

with $\widehat{v g}$ being an upwind numerical flux consistent to $v g$,

$$\widehat{v g} := \begin{cases} v g^-, & \text{if } v > 0 \\ v g^+, & \text{if } v < 0 \end{cases} = v \{g\} - \frac{|v|}{2} [g]. \quad (2.8)$$

Both $\langle \widehat{v g} \rangle$ and $\widehat{\rho}$ in (2.6a) and (2.6c) are also numerical fluxes, and they are consistent to the physical fluxes $\langle v g \rangle$ and ρ . In this paper, the following choices are considered:

$$\text{alternating: } \langle \widehat{v g} \rangle = \langle v g \rangle^-, \widehat{\rho} = \rho^+ \text{ (left-right); or } \langle \widehat{v g} \rangle = \langle v g \rangle^+, \widehat{\rho} = \rho^- \text{ (right-left),} \quad (2.9a)$$

$$\text{central: } \langle \widehat{v g} \rangle = \{ \langle v g \rangle \}, \widehat{\rho} = \{ \rho \}. \quad (2.9b)$$

Remark 2.1. The spatial discretization given above as proposed in [10] differs from the methods proposed in [19] in several aspects. First of all, the spatial discretization in [19] is first order and of finite difference type; while the method considered here is of finite element type, which is known to be more compact to achieve high order accuracy in a systematic manner. Moreover, our methods are based on one set of computational mesh and that in [19] is on staggered meshes. In fact, DG spatial discretizations can also be formulated on staggered meshes as in [19]. This, on one hand, saves one from using numerical fluxes $\langle \widehat{v g} \rangle$ and $\widehat{\rho}$ at grid points (see (2.6a), (2.6c), (2.9)), and on the other hand, the resulting method only has suboptimal k -th order accuracy when the P^k polynomial space is used with odd k (this is not reported yet observed numerically, and it is similar to our method with central fluxes for $\langle \widehat{v g} \rangle$ and $\widehat{\rho}$), while better spatial accuracy can be achieved both theoretically and numerically by suitably designing $\langle \widehat{v g} \rangle$ and $\widehat{\rho}$ within the one-mesh framework.

The semi-discrete method in (2.5) will be further coupled with globally stiffly accurate IMEX Runge-Kutta methods in time [3]. Such temporal discretizations are employed to deal with the stiffness of (2.3) when ε is small, and to ensure the correct asymptotic property of the scheme as $\varepsilon \rightarrow 0$. Below we will give the fully discrete scheme with a *first* order globally stiffly accurate IMEX Runge-Kutta method in time, termed as DG-IMEX1, which will be analyzed in the present paper. Methods with higher order temporal accuracy are formulated for the telegraph equation in [10], where the stability and accuracy are demonstrated numerically. Such high order temporal discretizations can also be defined for the one-group transport equation. The analysis for the fully discrete methods with higher than first order temporal accuracy is expected to be much more involved (see e.g. for such analysis in [28, 29] when the temporal discretizations are explicit) and it will be investigated in the next stage of our project.

Given $\rho_h^n(\cdot), g_h^n(\cdot, v) \in U_h^k$ that approximate the solution ρ and g at $t = t^n$, we look for $\rho_h^{n+1}(\cdot), g_h^{n+1}(\cdot, v) \in U_h^k$, such that $\forall \phi, \psi \in U_h^k$,

$$\left(\frac{\rho_h^{n+1} - \rho_h^n}{\Delta t}, \phi \right) + a_h(g_h^n, \phi) = 0, \quad (2.10a)$$

$$\left(\frac{g_h^{n+1} - g_h^n}{\Delta t}, \psi \right) + \frac{1}{\varepsilon} b_{h,v}(g_h^n, \psi) - \frac{v}{\varepsilon^2} d_h(\rho_h^{n+1}, \psi) = -\frac{1}{\varepsilon^2} (g_h^{n+1}, \psi). \quad (2.10b)$$

The most stiff terms, in both the convective and collisional terms, with a scale of $\frac{1}{\varepsilon^2}$ are treated implicitly here.

Remark 2.2. The implicit-explicit strategy used in our temporal discretization is different from that in [19]. We consider it more natural to treat implicitly the most stiff terms with the scale of $\frac{1}{\varepsilon^2}$ in the micro-macro reformulation, and there is no essential change in the computational complexity. Specifically, one can solve (2.10a) first for ρ_h^{n+1} , then (2.10b) for g_h^{n+1} from a block-diagonal system, indicating our implicit-explicit strategy results in comparable computational complexity as that in [19]. Moreover, this implicit-explicit strategy is especially important when it is combined with high order globally stiffly accurate temporal discretizations in [10]: it not only ensures the limiting schemes as $\varepsilon \rightarrow 0$ to be consistent high order discretizations for the limiting equations, but also preserves the limiting equilibrium on the discrete level, in the sense that

$$v d_h(\rho_h^*, \psi) = (g_h^*, \psi), \quad \psi \in U_h^k, \quad (2.11)$$

is satisfied by (ρ_h^*, g_h^*) , the approximating solution from any of the internal stages over one time step or at any discrete time t^n in the limit of $\varepsilon \rightarrow 0$ (see also Section 3.2 in [10]). The difference in the implicit-explicit strategy calls for a non-conventional definition of the discrete energy in Theorem 3.3. In Remark 3.5 of Section 3, there is also some discussion about how the theoretical results here and in [19, 20] are related.

3 Theoretical Results: Stability, Error Estimates, and Rigorous Asymptotic Analysis

In this section, stability and error estimates will be established for DG-IMEX1 method in (2.10) for both the telegraph equation and the one-group transport equation in slab geometry based on their micro-macro formulation (2.3). One will see that uniform stability result is obtained with respect to ε . In terms of accuracy, the method is first order in time for any given ε . The accuracy in space is higher when polynomials of higher degree are used for spatial approximations. When the exact solutions are bounded uniformly with respect to ε in certain norms, the error estimates are also uniform in ε . In addition, a rigorous asymptotic analysis is presented for the proposed methods when $\varepsilon \rightarrow 0$.

Without loss of generality, the mesh is assumed to be uniform with $h = h_i, \forall i$. Our results can be extended to general meshes when $\frac{\max_i h_i}{\min_i h_i}$ is uniformly bounded during the mesh refinement. Moreover, with very little change, our analysis can be established for the one-group transport equation in slab geometry with the more general scattering operator as in Section 2 (some constants in the results will also depend on the bounds s_m and s_M of the kernel $s(\cdot, \cdot)$). For simplicity, we will not present the analysis for the general case. The analysis will be based on the following norms,

$$\|\phi\| = \|\phi\|_{L^2(\Omega_x)}, \quad \|\phi\| = ((\|\phi\|^2))^{1/2}.$$

In the estimates, two standard inverse inequalities will be used [7]. There exist constants $C_{\text{inv}}, \hat{C}_{\text{inv}}$, such that for any $w \in P^k([a, b])$,

$$|w(y)|^2(b-a) \leq C_{\text{inv}} \int_a^b w(x)^2 dx, \quad \text{with } y = a, \text{ or } b \quad (3.12a)$$

$$(b-a)^2 \int_a^b |w_x(x)|^2 dx \leq \hat{C}_{\text{inv}} \int_a^b w(x)^2 dx. \quad (3.12b)$$

The constants $C_{\text{inv}}, \hat{C}_{\text{inv}}$ are independent of a and b , and they depend on k (see [27, 25] for the explicit expression of k -dependence). Some other basic inequalities, such as Young's inequality $xy \leq \frac{x^2}{2\zeta} + \frac{\zeta y^2}{2}$ ($\zeta > 0$), are used without being pointed out. We also denote

$$\alpha_1 = (\|v\|_\infty^2 + \langle v^2 \rangle) \hat{C}_{\text{inv}}, \quad \alpha_2 = 2(\|v\|_\infty + \langle |v| \rangle) C_{\text{inv}}, \quad \alpha_3 = 2\|v\|_\infty C_{\text{inv}}, \quad (3.13)$$

for later use. Below are some orthogonal projections onto U_h^k utilized in our analysis.

- L^2 projection π_h : $\pi_h w \in U_h^k$, such that

$$\int_{I_i} (\pi_h w - w) v dx = 0, \quad \forall v \in P^k(I_i), \forall i.$$

- Gauss-Radau projection π_h^- : $\pi_h^- w \in U_h^k$, such that

$$\int_{I_i} (\pi_h^- w - w) v dx = 0, \quad \forall v \in P^{k-1}(I_i), \forall i$$

$$\text{and } (\pi_h^- w)_{i+\frac{1}{2}}^- = w(x_{i+\frac{1}{2}}^-), \forall i.$$

- Gauss-Radau projection π_h^+ : $\pi_h^+ w \in U_h^k$, such that

$$\int_{I_i} (\pi_h^+ w - w) v dx = 0, \quad \forall v \in P^{k-1}(I_i), \forall i$$

$$\text{and } (\pi_h^+ w)_{i-\frac{1}{2}}^+ = w(x_{i-\frac{1}{2}}^+), \forall i.$$

These projections are commonly used in theoretical analysis of DG methods, and they have the following properties which can be easily established [7],

$$\|w - \Pi_h w\|^2 + h \sum_i ((w - \Pi_h w)_{i-\frac{1}{2}}^\pm)^2 \leq Ch^{2k+2} \|w\|_{H^{k+1}(\Omega_x)}^2. \quad (3.14)$$

Here Π_h is any of π_h, π_h^-, π_h^+ , and the constant C depends only on k .

At $t = 0$, we initialize the methods through L^2 projection in space, that is, we take $\rho_h^0 = \pi_h \rho|_{t=0}$ and $g_h^0 = \pi_h g|_{t=0}$. It can be seen easily that

$$\langle g_h^0 \rangle = \pi_h \langle g \rangle|_{t=0} = 0. \quad (3.15)$$

Other types of initialization, for example $\rho_h^0 = \Pi_\rho \rho|_{t=0}$ and $g_h^0 = \Pi_g g|_{t=0}$ with Π_ρ and Π_g specified in Section 3.2, can be considered without any essential change to the results presented in this section.

3.1 Stability analysis

To prove stability, we first present two lemmas. In particular, Lemma 3.1 is a discrete analogue of the property $\langle g \rangle = 0$ for the exact solution. It plays an important role in both stability analysis and error estimates.

Lemma 3.1. The numerical solution g_h^n has the following property

$$\langle g_h^n \rangle = 0, \quad \forall n. \quad (3.16)$$

Proof. Take $\forall \psi \in U_h^k$ in (2.10b), integrate over v , with $\langle v \rangle = 0$, one has

$$\left\langle \left(\frac{g_h^{n+1} - g_h^n}{\Delta t}, \psi \right) \right\rangle = -\frac{1}{\varepsilon^2} \langle (g_h^{n+1}, \psi) \rangle.$$

This can be further organized into

$$\langle (g_h^{n+1}, \psi) \rangle = \frac{\varepsilon^2}{\varepsilon^2 + \Delta t} \langle (g_h^n, \psi) \rangle, \quad \forall \psi \in U_h^k. \quad (3.17)$$

By taking $\psi = \langle g_h^{n+1} \rangle - \frac{\varepsilon^2}{\varepsilon^2 + \Delta t} \langle g_h^n \rangle \in U_h^k$, in addition to the fact $\langle g_h^0 \rangle = 0$ in (3.15), one can conclude $\langle g_h^{n+1} \rangle = \frac{\varepsilon^2}{\varepsilon^2 + \Delta t} \langle g_h^n \rangle$ hence (3.16). \square

Lemma 3.2. Let n and m be any non-negative integer indices. For any $\phi^n(\cdot), \phi^{n+1}(\cdot), \psi^m(\cdot, v), \psi^{m+1}(\cdot, v) \in U_h^k$ satisfying $\langle \psi^m \rangle = \langle \psi^{m+1} \rangle = 0$, denote

$$\begin{aligned} \Xi_\varepsilon(\phi^{n+1}, \phi^n, \psi^{m+1}, \psi^m) &:= \left(\frac{\phi^{n+1} - \phi^n}{\Delta t}, \phi^{n+1} \right) + a_h(\psi^{m+1}, \phi^{n+1}) \\ &+ \left\langle \left(\frac{\psi^{m+1} - \psi^m}{\Delta t}, \varepsilon^2 \psi^{m+1} \right) + \frac{1}{\varepsilon} b_{h,v}(\psi^m, \varepsilon^2 \psi^{m+1}) - \frac{v}{\varepsilon^2} d_h(\phi^n, \varepsilon^2 \psi^{m+1}) \right\rangle, \end{aligned} \quad (3.18)$$

the following estimates then hold with any of the numerical flux in (2.9),

$$\begin{aligned} \Xi_\varepsilon(\phi^{n+1}, \phi^n, \psi^{m+1}, \psi^m) &\geq \frac{1}{2\Delta t} ((\|\phi^{n+1}\|^2 + \varepsilon^2 \|\psi^{m+1}\|^2) - (\|\phi^n\|^2 + \varepsilon^2 \|\psi^m\|^2)) \\ &+ \begin{cases} (\varepsilon - \alpha_2 \frac{\Delta t}{h}) \langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \rangle - \alpha_1 \frac{\Delta t}{h^2} \|\psi^{m+1}\|^2, & \text{for } k \geq 1 \\ (\varepsilon - \frac{\alpha_2 \Delta t}{2h}) \langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \rangle, & \text{for } k = 0, \end{cases} \end{aligned} \quad (3.19)$$

with $\alpha_i, i = 1, 2, 3$ defined in (3.13).

Proof. Note that

$$\begin{aligned} \Xi_\varepsilon(\phi^{n+1}, \phi^n, \psi^{m+1}, \psi^m) &= \frac{1}{2\Delta t} (\|\phi^{n+1}\|^2 - \|\phi^n\|^2 + \|\phi^{n+1} - \phi^n\|^2) + \frac{\varepsilon^2}{2\Delta t} (\|\psi^{m+1}\|^2 - \|\psi^m\|^2 + \|\psi^{m+1} - \psi^m\|^2) \\ &+ a_h(\psi^{m+1}, \phi^{n+1}) + \varepsilon \langle b_{h,v}(\psi^m, \psi^{m+1}) \rangle - \langle v d_h(\phi^{n+1}, \psi^{m+1}) \rangle + \langle v d_h(\phi^{n+1} - \phi^n, \psi^{m+1}) \rangle. \end{aligned} \quad (3.20)$$

Based on the definitions of the bilinear forms $a_h(\cdot, \cdot)$ and $d_h(\cdot, \cdot)$, one has

$$\begin{aligned} &a_h(\psi^{m+1}, \phi^{n+1}) - \langle v d_h(\phi^{n+1}, \psi^{m+1}) \rangle \\ &= - \sum_i \int_{I_i} \partial_x (\langle v \psi^{m+1} \rangle \phi^{n+1}) dx - \sum_i \langle \widehat{v \psi^{m+1}} \rangle_{i-\frac{1}{2}} [\phi^{n+1}]_{i-\frac{1}{2}} - \sum_i \widehat{\phi_i^{n+1}} [\langle v \psi^{m+1} \rangle]_{i-\frac{1}{2}} \\ &= \sum_i \left([\langle v \psi^{m+1} \rangle \phi^{n+1}] - \langle \widehat{v \psi^{m+1}} \rangle [\phi^{n+1}] - \widehat{\phi_i^{n+1}} [\langle v \psi^{m+1} \rangle] \right)_{i-\frac{1}{2}} = 0. \end{aligned} \quad (3.21)$$

The last equality can be verified directly with the definition of central and alternating fluxes. In addition, with $\langle \psi^{m+1} \rangle = 0$ and equation (2.7) and the upwind flux (2.8),

$$\begin{aligned} \langle b_{h,v}(\psi^m, \psi^{m+1}) \rangle &= \langle (\mathcal{D}_h(\psi^m; v) - \langle \mathcal{D}_h(\psi^m; v) \rangle), \psi^{m+1} \rangle \\ &= \langle (\mathcal{D}_h(\psi^m; v), \psi^{m+1}) \rangle - \langle (\mathcal{D}_h(\psi^m; v)), \langle \psi^{m+1} \rangle \rangle = \langle (\mathcal{D}_h(\psi^m; v), \psi^{m+1}) \rangle \\ &= \langle (\mathcal{D}_h(\psi^{m+1}; v), \psi^{m+1}) \rangle + \langle (\mathcal{D}_h(\psi^m - \psi^{m+1}; v), \psi^{m+1}) \rangle \\ &= - \sum_i \left(\left\langle \int_{I_i} v \psi^{m+1} \partial_x \psi^{m+1} dx + (\widehat{v \psi^{m+1}})_{i-\frac{1}{2}} [\psi^{m+1}]_{i-\frac{1}{2}} \right\rangle \right) + \langle (\mathcal{D}_h(\psi^m - \psi^{m+1}; v), \psi^{m+1}) \rangle \\ &= \langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \rangle + \langle (\mathcal{D}_h(\psi^m - \psi^{m+1}; v), \psi^{m+1}) \rangle. \end{aligned} \quad (3.22)$$

Next we want to estimate $\langle (\mathcal{D}_h(\psi^m - \psi^{m+1}; v), \psi^{m+1}) \rangle = \Lambda_1 + \Lambda_2$ in (3.22), where

$$\Lambda_1 = \left\langle \sum_i \int_{I_i} v (\psi^{m+1} - \psi^m) \partial_x \psi^{m+1} dx \right\rangle, \quad \Lambda_2 = \left\langle \sum_i (v (\widehat{\psi^{m+1}} - \psi^m))_{i-\frac{1}{2}} [\psi^{m+1}]_{i-\frac{1}{2}} \right\rangle.$$

These two terms can be bounded as follows.

$$|\Lambda_1| \leq \theta_1 \|\psi^{m+1} - \psi^m\|^2 + \frac{1}{4\theta_1} \left\langle \int_{\Omega_x} (v \partial_x \psi^{m+1})^2 dx \right\rangle, \quad (3.23)$$

$$\begin{aligned}
|\Lambda_2| &= \left| \left\langle \sum_i v(\psi^{m+1} - \psi^m)_{i-\frac{1}{2}}^- [\psi^{m+1}]_{i-\frac{1}{2}} \right\rangle_+ + \left\langle \sum_i v(\psi^{m+1} - \psi^m)_{i-\frac{1}{2}}^+ [\psi^{m+1}]_{i-\frac{1}{2}} \right\rangle_- \right| \\
&\leq \frac{\eta_1}{C_{\text{inv}}} \left\langle \sum_i h((\psi^{m+1} - \psi^m)_{i-\frac{1}{2}}^-)^2 \right\rangle_+ + \frac{C_{\text{inv}}}{4\eta_1} \left\langle \sum_i h^{-1}(v[\psi^{m+1}]_{i-\frac{1}{2}})^2 \right\rangle_+ \\
&\quad + \frac{\eta_1}{C_{\text{inv}}} \left\langle \sum_i h((\psi^{m+1} - \psi^m)_{i-\frac{1}{2}}^+)^2 \right\rangle_- + \frac{C_{\text{inv}}}{4\eta_1} \left\langle \sum_i h^{-1}(v[\psi^{m+1}]_{i-\frac{1}{2}})^2 \right\rangle_- \\
&\leq \eta_1 \|\psi^{m+1} - \psi^m\|^2 + \frac{C_{\text{inv}}}{4\eta_1} \sum_i h^{-1} \langle (v[\psi^{m+1}]_{i-\frac{1}{2}})^2 \rangle. \tag{3.24}
\end{aligned}$$

Here the inverse inequality in (3.12a) is used, and $\langle \cdot \rangle_+$ and $\langle \cdot \rangle_-$ are integrals with respect to the positive and negative part of v , respectively. What we also need to estimate is $\langle v d_h(\phi^{n+1} - \phi^n, \psi^{m+1}) \rangle$.

$$\begin{aligned}
&|\langle v d_h(\phi^{n+1} - \phi^n, \psi^{m+1}) \rangle| \\
&\leq \left| \sum_i \int_{I_i} (\phi^{n+1} - \phi^n) \partial_x \langle v \psi^{m+1} \rangle dx \right| + \left| \sum_i \left((\widehat{\phi^{n+1}} - \phi^n) \langle v[\psi^{m+1}] \rangle \right)_{i-\frac{1}{2}} \right| \\
&\leq \theta_2 \|\phi^{n+1} - \phi^n\|^2 + \frac{1}{4\theta_2} \int_{\Omega_x} (\partial_x \langle v \psi^{m+1} \rangle)^2 dx + \frac{\eta_2}{C_{\text{inv}}} \sum_i h(\widehat{\phi^{n+1}} - \phi^n)_{i-\frac{1}{2}}^2 + \frac{C_{\text{inv}}}{4\eta_2} \sum_i h^{-1} \langle v[\psi^{m+1}]_{i-\frac{1}{2}} \rangle^2 \\
&\leq \theta_2 \|\phi^{n+1} - \phi^n\|^2 + \frac{1}{4\theta_2} \int_{\Omega_x} (\partial_x \langle v \psi^{m+1} \rangle)^2 dx + \eta_2 \|\phi^{n+1} - \phi^n\|^2 + \frac{C_{\text{inv}}}{4\eta_2} \sum_i h^{-1} \langle v[\psi^{m+1}]_{i-\frac{1}{2}} \rangle^2 \\
&= (\theta_2 + \eta_2) \|\phi^{n+1} - \phi^n\|^2 + \frac{1}{4\theta_2} \int_{\Omega_x} (\partial_x \langle v \psi^{m+1} \rangle)^2 dx + \frac{C_{\text{inv}}}{4\eta_2} \sum_i h^{-1} \langle v[\psi^{m+1}]_{i-\frac{1}{2}} \rangle^2. \tag{3.25}
\end{aligned}$$

Up to now, θ_i, η_i with $i = 1, 2$ are arbitrary positive constants. By further applying inverse inequalities in (3.12b) and the following simple estimates,

$$\begin{aligned}
\left\langle \int_{\Omega_x} (v \partial_x \psi^{m+1})^2 dx \right\rangle &\leq \|v\|_\infty^2 \left\langle \int_{\Omega_x} (\partial_x \psi^{m+1})^2 dx \right\rangle \leq \frac{\hat{C}_{\text{inv}}}{h^2} \|v\|_\infty^2 \|\psi^{m+1}\|^2, \\
\int_{\Omega_x} (\partial_x \langle v \psi^{m+1} \rangle)^2 dx &= \int_{\Omega_x} \langle v \partial_x \psi^{m+1} \rangle^2 dx \leq \int_{\Omega_x} \langle v^2 \rangle \langle (\partial_x \psi^{m+1})^2 \rangle dx \\
&= \langle v^2 \rangle \left\langle \int_{\Omega_x} (\partial_x \psi^{m+1})^2 dx \right\rangle \leq \frac{\hat{C}_{\text{inv}}}{h^2} \langle v^2 \rangle \|\psi^{m+1}\|^2, \\
\sum_i \langle (v[\psi^{m+1}]_{i-\frac{1}{2}})^2 \rangle &\leq 2\|v\|_\infty \left\langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \right\rangle, \\
\sum_i \langle v[\psi^{m+1}]_{i-\frac{1}{2}} \rangle^2 &\leq \sum_i \langle |v| \rangle \langle |v| ([\psi^{m+1}]_{i-\frac{1}{2}})^2 \rangle = 2\langle |v| \rangle \left\langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \right\rangle,
\end{aligned}$$

we have

$$\begin{aligned}
|\langle (\mathcal{D}_h(\psi^m - \psi^{m+1}; v), \psi^{m+1}) \rangle| &\leq (\theta_1 + \eta_1) \|\psi^{m+1} - \psi^m\|^2 + \frac{\hat{C}_{\text{inv}} \|v\|_\infty^2}{4\theta_1 h^2} \|\psi^{m+1}\|^2 \\
&\quad + \frac{C_{\text{inv}} \|v\|_\infty}{2\eta_1 h} \left\langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \right\rangle, \tag{3.26}
\end{aligned}$$

$$\begin{aligned}
|\langle v d_h(\phi^{n+1} - \phi^n, \psi^{m+1}) \rangle| &\leq (\theta_2 + \eta_2) \|\phi^{n+1} - \phi^n\|^2 + \frac{\hat{C}_{\text{inv}} \langle v^2 \rangle}{4\theta_2 h^2} \|\psi^{m+1}\|^2 \\
&\quad + \frac{C_{\text{inv}} \langle |v| \rangle}{2\eta_2 h} \left\langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \right\rangle. \tag{3.27}
\end{aligned}$$

Estimates in (3.20)-(3.22) and (3.26)-(3.27) are now assembled together,

$$\begin{aligned}
\Xi_\varepsilon(\phi^{n+1}, \phi^n, \psi^{m+1}, \psi^m) &\geq \frac{1}{2\Delta t} ((\|\phi^{n+1}\|^2 + \varepsilon^2\|\psi^{m+1}\|^2) - (\|\phi^n\|^2 + \varepsilon^2\|\psi^m\|^2)) \\
&+ \left(\frac{1}{2\Delta t} - (\theta_2 + \eta_2)\right) \|\phi^{n+1} - \phi^n\|^2 + \left(\frac{\varepsilon^2}{2\Delta t} - \varepsilon(\theta_1 + \eta_1)\right) \|\psi^{m+1} - \psi^m\|^2 \\
&+ \left(\varepsilon - \frac{\varepsilon C_{\text{inv}}\|v\|_\infty}{2\eta_1 h} - \frac{C_{\text{inv}}\langle|v|\rangle}{2\eta_2 h}\right) \left\langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \right\rangle \\
&- \left(\frac{\varepsilon \hat{C}_{\text{inv}}\|v\|_\infty^2}{4\theta_1 h^2} + \frac{\hat{C}_{\text{inv}}\langle v^2 \rangle}{4\theta_2 h^2}\right) \|\psi^{m+1}\|^2.
\end{aligned} \tag{3.28}$$

One can conclude the lower bound of Ξ_ε in (3.19) by taking $\theta_1 = \eta_1 = \frac{\varepsilon}{4\Delta t}$, $\theta_2 = \eta_2 = \frac{1}{4\Delta t}$ in (3.28). When $k = 0$, $\partial_x \psi^{m+1} = 0$ and all terms involving θ_1 and θ_2 are no longer needed in the analysis. The estimate in (3.28) turns to

$$\begin{aligned}
\Xi_\varepsilon(\phi^{n+1}, \phi^n, \psi^{m+1}, \psi^m) &\geq \frac{1}{2\Delta t} ((\|\phi^{n+1}\|^2 + \varepsilon^2\|\psi^{m+1}\|^2) - (\|\phi^n\|^2 + \varepsilon^2\|\psi^m\|^2)) \\
&+ \left(\frac{1}{2\Delta t} - \eta_2\right) \|\phi^{n+1} - \phi^n\|^2 + \left(\frac{\varepsilon^2}{2\Delta t} - \varepsilon\eta_1\right) \|\psi^{m+1} - \psi^m\|^2 \\
&+ \left(\varepsilon - \frac{\varepsilon C_{\text{inv}}\|v\|_\infty}{2\eta_1 h} - \frac{C_{\text{inv}}\langle|v|\rangle}{2\eta_2 h}\right) \left\langle \frac{|v|}{2} \sum_i [\psi^{m+1}]_{i-\frac{1}{2}}^2 \right\rangle.
\end{aligned} \tag{3.29}$$

In particular, with $\eta_1 = \frac{\varepsilon}{2\Delta t}$ and $\eta_2 = \frac{1}{2\Delta t}$ in (3.29), we conclude the estimate for $k = 0$. \square

Theorem 3.3 (Stability of DG-IMEX1). *When the DG-IMEX1 method (2.10) is applied to the kinetic transport equation (2.1) in its micro-macro decomposition formulation (2.3), the following stability result holds for the numerical solution,*

$$\|\rho_h^{n+1}\|^2 + \varepsilon^2\|g_h^n\|^2 \leq \|\rho_h^n\|^2 + \varepsilon^2\|g_h^{n-1}\|^2, \forall n \tag{3.30}$$

under the condition

$$\Delta t \leq \Delta t_{\text{stab}} = \begin{cases} \frac{h}{\alpha_1 + \alpha_2 \alpha_3} (h + \min(\varepsilon, \frac{\alpha_2 h}{\alpha_1}) \alpha_3), & \text{for } k \geq 1, \\ \frac{2h}{\alpha_2 \alpha_3} (h + \alpha_3 \varepsilon), & \text{for } k = 0. \end{cases} \tag{3.31}$$

Here $\alpha_i, i = 1, 2, 3$ are defined in (3.13).

Proof. Take $\phi = \rho_h^{n+1}$ in (2.10a). Additionally take $\psi = \varepsilon^2 g_h^{n+1}$ in (2.10b), integrate the equation over v , and shift the index n to $n - 1$. This leads to

$$\left(\frac{\rho_h^{n+1} - \rho_h^n}{\Delta t}, \rho_h^{n+1}\right) + a_h(g_h^n, \rho_h^{n+1}) = 0, \tag{3.32a}$$

$$\left\langle \left(\frac{g_h^n - g_h^{n-1}}{\Delta t}, \varepsilon^2 g_h^n\right) + \frac{1}{\varepsilon} b_{h,v}(g_h^{n-1}, \varepsilon^2 g_h^n) - \frac{v}{\varepsilon^2} d_h(\rho_h^n, \varepsilon^2 g_h^n) \right\rangle = -\left\langle \frac{1}{\varepsilon^2} (g_h^n, \varepsilon^2 g_h^n) \right\rangle. \tag{3.32b}$$

Now we sum up equations (3.32a)-(3.32b), denote the left side of the resulting equation as LHS , and get

$$LHS = -\|g_h^n\|^2. \tag{3.33}$$

Let's first consider $k \geq 1$. For any integer index $n \geq 1$, by applying Lemma 3.2 with $\phi = \rho_h$, $\psi = g_h$, and $m = n - 1$, one gets

$$\begin{aligned}
LHS = \Xi_\varepsilon(\rho_h^{n+1}, \rho_h^n, g_h^n, g_h^{n-1}) &\geq \frac{1}{2\Delta t} ((\|\rho_h^{n+1}\|^2 + \varepsilon^2\|g_h^n\|^2) - (\|\rho_h^n\|^2 + \varepsilon^2\|g_h^{n-1}\|^2)) \\
&+ \left(\varepsilon - \alpha_2 \frac{\Delta t}{h}\right) \left\langle \frac{|v|}{2} \sum_i [g_h^n]_{i-\frac{1}{2}}^2 \right\rangle - \alpha_1 \frac{\Delta t}{h^2} \|g_h^n\|^2.
\end{aligned} \tag{3.34}$$

Note that

$$\left\langle \frac{|v|}{2} \sum_i [g_h^n]_{i-\frac{1}{2}}^2 \right\rangle \leq \|v\|_\infty \left\langle \sum_i (g_{h,i-\frac{1}{2}}^{n,+})^2 + (g_{h,i-\frac{1}{2}}^{n,-})^2 \right\rangle \leq \frac{2\|v\|_\infty C_{\text{inv}}}{h} \|g_h^n\|^2 = \frac{\alpha_3}{h} \|g_h^n\|^2. \tag{3.35}$$

Combining this estimate with (3.33)-(3.34), one has

$$\frac{1}{2\Delta t} (\|\rho_h^{n+1}\|^2 + \varepsilon^2 \|g_h^n\|^2) - (\|\rho_h^n\|^2 + \varepsilon^2 \|g_h^{n-1}\|^2) \leq \gamma \|g_h^n\|^2, \quad (3.36)$$

where

$$\gamma = \left(-1 + \alpha_1 \frac{\Delta t}{h^2}\right) + \max\left(-\varepsilon + \alpha_2 \frac{\Delta t}{h}, 0\right) \frac{\alpha_3}{h}. \quad (3.37)$$

The stability result (3.30) can be obtained now as long as $\gamma \leq 0$. Equivalently, this requires

$$-\varepsilon + \alpha_2 \frac{\Delta t}{h} \leq 0 \quad \text{and} \quad -1 + \alpha_1 \frac{\Delta t}{h^2} \leq 0, \quad (3.38)$$

or

$$-\varepsilon + \alpha_2 \frac{\Delta t}{h} > 0 \quad \text{and} \quad -1 + \alpha_1 \frac{\Delta t}{h^2} + (-\varepsilon + \alpha_2 \frac{\Delta t}{h}) \frac{\alpha_3}{h} \leq 0. \quad (3.39)$$

These conditions can be reformulated into

Case 1: when $\frac{\varepsilon h}{\alpha_2} \leq \frac{h^2}{\alpha_1}$, that is $\varepsilon \leq \frac{\alpha_2}{\alpha_1} h$, then $\Delta t \leq \frac{h^2 + \alpha_3 \varepsilon h}{\alpha_1 + \alpha_2 \alpha_3}$.

Case 2: when $\varepsilon > \frac{\alpha_2}{\alpha_1} h$, then $\Delta t \leq \frac{h^2}{\alpha_1}$.

Conditions in both cases can be compactly written into

$$\Delta t \leq \frac{h}{\alpha_1 + \alpha_2 \alpha_3} (h + \min(\varepsilon, \frac{\alpha_2 h}{\alpha_1}) \alpha_3). \quad (3.40)$$

Finally, we consider $k = 0$. By following similar analysis as above using the corresponding result in Lemma 3.2, one will have (3.36) where

$$\gamma = -1 + \max\left(-\varepsilon + \frac{\alpha_2}{2} \frac{\Delta t}{h}, 0\right) \frac{\alpha_3}{h}. \quad (3.41)$$

To conclude the stability result (3.30), it is required to have $\gamma \leq 0$ which is equivalent to $-1 + (-\varepsilon + \frac{\alpha_2}{2} \frac{\Delta t}{h}) \frac{\alpha_3}{h} \leq 0$. This condition can be further simplified into $\Delta t \leq \frac{2h}{\alpha_2 \alpha_3} (h + \alpha_3 \varepsilon)$. This completes the proof. \square

Note that the stability is measured in a non-conventional way in that the discrete energy in (3.30) at the n -th step consists of the L^2 norm of ρ_h^n , and the L^2 norm of g_h^{n-1} . This is due to the implicit-explicit strategy used in the numerical formulation. One can refer to [19] [20] for a different implicit-explicit strategy used in a first order scheme, hence the stability analysis with different discrete energy. By working out the expression of $\alpha_i, i = 1, 2, 3$ in (3.13) for the specific equations in Section 2, we further have the following remark.

Remark 3.4. The condition (3.31) is

$$\Delta t \leq \Delta t_{stab} = \begin{cases} \frac{1}{2C_{inv} + 8C_{inv}^2} (h + 2C_{inv} \min(\varepsilon, \frac{2C_{inv}}{C_{inv}} h)) h, & \text{for } k \geq 1, \\ \frac{h}{4C_{inv}^2} (h + 2C_{inv} \varepsilon), & \text{for } k = 0. \end{cases} \quad (3.42)$$

for the telegraph equation, and it is

$$\Delta t \leq \Delta t_{stab} = \begin{cases} \frac{1}{\frac{4}{3}C_{inv} + 6C_{inv}^2} (h + 2C_{inv} \min(\varepsilon, \frac{9}{4} \frac{C_{inv}}{C_{inv}} h)) h, & \text{for } k \geq 1, \\ \frac{h}{3C_{inv}^2} (h + 2C_{inv} \varepsilon), & \text{for } k = 0. \end{cases} \quad (3.43)$$

for the one-group transport equation in slab geometry.

Remark 3.5. • The stability condition in (3.31) for the DG-IMEX1 is established *uniformly* with respect to ε for any given integer $k \geq 0$.

- When $k = 0$, one has $C_{inv} = 1$. (The actual values of C_{inv} and \hat{C}_{inv} for $k \geq 1$ can be found in [27, 25].) The stability condition in (3.42) becomes $\Delta t \leq \frac{1}{4} h^2 + \frac{1}{2} \varepsilon h$, and the one in (3.43) is $\Delta t \leq \frac{1}{3} h^2 + \frac{2}{3} \varepsilon h$. These results are the same as that of the first order finite difference method introduced in [19] on staggered grids which employs a different implicit-explicit strategy in discretization (see [19] for the telegraph equation and [20] for the one-group transport equation in slab geometry). Moreover, when ε is small with the equation in the diffusive regime, $\Delta t = O(h^2)$; when ε is large with the equation in the convective regime, $\Delta t = O(\varepsilon h)$. The timestep restrictions in both regimes are standard for explicit schemes.

- For $k \geq 1$, the restriction on the time step is $\Delta t = O(h^2)$, which is reasonable for $\varepsilon \ll 1$, yet not the most desired condition for $\varepsilon = O(1)$ in the convective regime. Similar as in [29] for analyzing DG methods with explicit Runge-Kutta time discretizations, we conjecture that higher order time discretizations will improve or overcome the restrictive condition on the time step when $\varepsilon = O(1)$. This will be left to our future investigation.
- For the telegraph equation, with its special discrete velocity space $v \in \{-1, 1\}$, one can verify that $(\mathbf{I} - \Pi)(v\partial_x g) = v\langle\partial_x g\rangle = v\partial_x\langle g\rangle$ holds. Since the exact solution satisfies $\langle g\rangle = 0$, it seems one does not need to include $(\mathbf{I} - \Pi)(v\partial_x g)$ in the micro-macro formulation (2.3). Numerically this means not to include $b_{h,v}$ term in (2.5) and in (2.10). For the resulting scheme, we can follow the similar analysis as in this subsection and obtain the stability result (3.30) under the condition

$$\Delta t \leq \Delta t_{stab} = \begin{cases} \frac{1}{C_{inv} + 4C_{inv}^2} h^2, & \text{for } k \geq 1, \\ \frac{1}{2C_{inv}^2} h^2 = \frac{1}{2} h^2, & \text{for } k = 0. \end{cases} \quad (3.44)$$

The time step constraint for stability is no longer reflecting the ‘‘multi-scale’’ aspect of the equation, and it is independent of ε and always $\Delta t = O(h^2)$, even in the convective regime with $\varepsilon = O(1)$ and when $k = 0$. This shows the importance of including the term $(\mathbf{I} - \Pi)(v\partial_x g)$ in the design of numerical methods.

3.2 Error estimates

In this subsection, error estimates are carried out for the proposed method (2.10) to solve the kinetic transport equation (2.1) in its micro-macro formulation (2.3) with smooth exact solutions at any given time $T : 0 < T < \infty$. Let ρ^n and g^n be the exact solution at time $t^n = n\Delta t, \forall n$. Let Π_ρ and Π_g denote two orthogonal projections onto U_h^k which will be specified later. Define the error function in ρ , $e_\rho^n = \rho^n - \rho_h^n = \xi_\rho^n - \eta_\rho^n$ where $\xi_\rho^n = \Pi_\rho \rho^n - \rho_h^n$ and $\eta_\rho^n = \Pi_\rho \rho^n - \rho^n$. Similarly, $e_g^n = g^n - g_h^n = \xi_g^n - \eta_g^n$ where $\xi_g^n = \Pi_g g^n - g_h^n$ and $\eta_g^n = \Pi_g g^n - g^n$. We also denote $\mathcal{E}_n = \|\xi_\rho^n\|^2 + \varepsilon^2 \|\xi_g^{n-1}\|^2$. Both η_ρ^n and η_g^n can be estimated in a standard way based on the definitions of Π_ρ , Π_g , and U_h^k (see the beginning of Section 3 regarding the property of projections), therefore the error estimates for the proposed methods boil down to the estimation of ξ_ρ^n and $\xi_g^n, \forall n$. Throughout this subsection, we use $C, C_* > 0$ to denote generic constants. Here C only depends on k ; C_* is independent of $h, \Delta t$ and n and depends on $k, T, \|v\|_\infty$ and some Sobolev norms of the exact solutions, more specifically, an upper bound of

$$\|\partial_{tt}\rho\|, \|\partial_{xt}\rho\|, \|\rho\|_{H^{k+1}(\Omega_x)}, \|\partial_t\rho\|_{H^{k+1}(\Omega_x)}, \|\partial_{tt}g\|, \langle \|g\|_{H^{k+1}(\Omega_x)}^2 \rangle, \langle \|\partial_tg\|_{H^{k+1}(\Omega_x)}^2 \rangle$$

over $t \in [0, T]$. Different occurrences of C, C_* could take different values. Standard notations for Sobolev spaces $H^{k+1}(\Omega_x)$ as well as their norms $\|\cdot\|_{H^{k+1}(\Omega_x)}$ are used in this paper [7].

Recall that at $t = 0$, the proposed methods are initialized through $\rho_h^0 = \pi_h \rho^0$ and $g_h^0 = \pi_h g^0$. With this and Lemma 3.1, we have $\langle \xi_g^n \rangle = \langle \Pi_g g^n \rangle - \langle g_h^n \rangle = \Pi_g \langle g^n \rangle = 0$, hence

$$\langle \xi_g^n \rangle = 0, \quad \forall n. \quad (3.45)$$

In next Theorem, we will state the main error estimate results, and their proofs will be given step by step in Sections 3.2.1-3.2.5.

Theorem 3.6 (Error estimate). *When the DG-IMEX1 method (2.10) is applied to the kinetic transport equation (2.1) in its micro-macro decomposition formulation (2.3), the following error estimates hold:*

(1) *with any of the alternating flux in (2.9),*

$$\begin{aligned} & \|\rho^n - \rho_h^n\|^2 + \varepsilon^2 \|g^{n-1} - g_h^{n-1}\|^2 \\ & \leq C_* \left((1 + \varepsilon^2) h^{2k+2} + \Delta t^2 + \frac{1}{1 - \sigma} ((1 + \varepsilon^4) \Delta t^2 + h^{2k+2} + \varepsilon h^{2k+1}) \right); \end{aligned} \quad (3.46)$$

(2) *with any of the central flux in (2.9),*

$$\begin{aligned} & \|\rho^n - \rho_h^n\|^2 + \varepsilon^2 \|g^{n-1} - g_h^{n-1}\|^2 \\ & \leq C_* \left(\varepsilon^2 h^{2k+2} + h^{2k} + \Delta t^2 + \frac{1}{1 - \sigma} ((1 + \varepsilon^4) \Delta t^2 + h^{2k} + \varepsilon h^{2k+1}) \right) \end{aligned} \quad (3.47)$$

Table 3.1: Spatial accuracy orders established by the error estimates.

| | alternating | central |
|----------------------|-------------------|---------|
| $\varepsilon = O(h)$ | $k + 1$ | k |
| $\varepsilon \gg h$ | $k + \frac{1}{2}$ | k |

for $n : n\Delta t \leq T$ under the condition $\Delta t \leq \sigma \Delta t_{stab}$ and $\Delta t < \frac{1}{2}$. Here σ is any constant in $(0, 1)$.

The error estimates are obtained as long as the time step is no larger than that required for numerical stability. We further summarize the established spatial accuracy orders in Table 3.1. Numerically, higher than theoretical convergence rates can be observed in some cases (see [10]).

Remark 3.7. What established here are a priori error estimates, and the constant C_* depends on exact solutions hence possibly on ε . For any test case where $\sup_\varepsilon C_* < \infty$, the error estimates in Theorem 3.6 hold uniformly with respect to ε .

3.2.1 Local truncation errors and error equations

Using the consistency of the DG spatial discretization, local truncation errors from the n -th step temporal discretization, denoted as $\tau_\rho^n(\cdot) \in U_h^k$ and $\tau_g^n(\cdot, v) \in U_h^k$, are defined as follows. For any $\phi(\cdot), \psi(\cdot, v) \in U_h^k$,

$$\left(\frac{\rho^{n+1} - \rho^n}{\Delta t}, \phi \right) + a_h(g^n, \phi) = (\tau_\rho^n, \phi), \quad (3.48a)$$

$$\left(\frac{g^{n+1} - g^n}{\Delta t}, \psi \right) + \frac{1}{\varepsilon} b_{h,v}(g^n, \psi) - \frac{v}{\varepsilon^2} d_h(\rho^{n+1}, \psi) = -\frac{1}{\varepsilon^2} (g^{n+1}, \psi) + (\tau_g^n, \psi). \quad (3.48b)$$

Lemma 3.8. The following estimates hold for the local truncation errors,

$$\|\tau_\rho^n\| \leq \frac{\Delta t}{\sqrt{3}} \max_{t \in [0, T]} \|\partial_{tt}\rho(\cdot, t)\|, \quad (3.49)$$

$$\|\varepsilon^2 \tau_g^n\| \leq \varepsilon^2 \frac{\Delta t}{\sqrt{3}} \max_{t \in [0, T]} \|\partial_{tt}g\| + \Delta t \max_{t \in [0, T]} (\|\partial_t g\| + \langle v^2 \rangle^{1/2} \|\partial_{xt}\rho\|). \quad (3.50)$$

Proof. For any $\phi \in U_h^k$, using the definition of the local truncation errors and integration by parts on each mesh element, we have

$$\begin{aligned} (\tau_\rho^n, \phi) &= \left(\frac{\rho^{n+1} - \rho^n}{\Delta t}, \phi \right) + a_h(g^n, \phi) = \left(\frac{\rho^{n+1} - \rho^n}{\Delta t} - \partial_t \rho^n, \phi \right) + (\partial_t \rho^n, \phi) + a_h(g^n, \phi) \\ &= \left(\frac{\rho^{n+1} - \rho^n}{\Delta t} - \partial_t \rho^n, \phi \right) + ((\partial_t \rho + \partial_x \langle vg \rangle)|_{t=t^n}, \phi) = \left(\frac{\rho^{n+1} - \rho^n}{\Delta t} - \partial_t \rho^n, \phi \right). \end{aligned}$$

The third equality is due to the consistency of the numerical flux, therefore the consistency of the spatial discretization. Now

$$\begin{aligned} \|\tau_\rho^n\| &= \max_{0 \neq \phi \in U_h^k} \frac{(\tau_\rho^n, \phi)}{\|\phi\|} \leq \left\| \frac{\rho^{n+1} - \rho^n}{\Delta t} - \partial_t \rho^n \right\| = \frac{1}{\Delta t} \left\| \int_{t^n}^{t^{n+1}} (t^{n+1} - t) \partial_{tt}\rho(x, t) dt \right\| \\ &\leq \frac{1}{\Delta t} \left(\int_{t^n}^{t^{n+1}} (t^{n+1} - t)^2 dt \int_{\Omega_x} \int_{t^n}^{t^{n+1}} (\partial_{tt}\rho(x, t))^2 dt dx \right)^{1/2} \leq \frac{\Delta t}{\sqrt{3}} \max_{t \in [0, T]} \|\partial_{tt}\rho(\cdot, t)\|. \end{aligned}$$

Similarly, for the truncation error in equation for g , $\forall \psi \in U_h^k$,

$$\begin{aligned}
(\varepsilon^2 \tau_g^n, \psi) &= \varepsilon^2 \left(\frac{g^{n+1} - g^n}{\Delta t}, \psi \right) + \varepsilon b_{h,v}(g^n, \psi) - v d_h(\rho^{n+1}, \psi) + (g^{n+1}, \psi) \\
&= \varepsilon^2 (\partial_t g^n, \psi) + \varepsilon b_{h,v}(g^n, \psi) - v d_h(\rho^n, \psi) + (g^n, \psi) \\
&\quad + \varepsilon^2 \left(\frac{g^{n+1} - g^n}{\Delta t} - \partial_t g^n, \psi \right) - v d_h(\rho^{n+1} - \rho^n, \psi) + (g^{n+1} - g^n, \psi) \\
&= \varepsilon^2 \left((\partial_t g + \frac{1}{\varepsilon^2} v \partial_x \rho + \frac{1}{\varepsilon} \{\mathbf{I} - \Pi\}(v \partial_x g) + \frac{1}{\varepsilon^2} g)|_{t=t^n}, \psi \right) \\
&\quad + \varepsilon^2 \left(\frac{g^{n+1} - g^n}{\Delta t} - \partial_t g^n, \psi \right) + (v \partial_x(\rho^{n+1} - \rho^n), \psi) + (g^{n+1} - g^n, \psi) \\
&= \varepsilon^2 \left(\frac{g^{n+1} - g^n}{\Delta t} - \partial_t g^n, \psi \right) + (v \partial_x(\rho^{n+1} - \rho^n), \psi) + (g^{n+1} - g^n, \psi).
\end{aligned}$$

Note $\|\varepsilon^2 \tau_g^n\| = \max_{0 \neq \psi \in U_h^k} \frac{(\varepsilon^2 \tau_g^n, \psi)}{\|\psi\|}$, then

$$\begin{aligned}
\|\varepsilon^2 \tau_g^n\| &\leq \|\varepsilon^2 (\frac{g^{n+1} - g^n}{\Delta t} - \partial_t g^n) + v \partial_x(\rho^{n+1} - \rho^n) + (g^{n+1} - g^n)\| \\
&\leq \varepsilon^2 \|\frac{g^{n+1} - g^n}{\Delta t} - \partial_t g^n\| + \|v \partial_x(\rho^{n+1} - \rho^n)\| + \|(g^{n+1} - g^n)\| \\
&\leq \varepsilon^2 \frac{\Delta t}{\sqrt{3}} \max_{t \in [0, T]} \|\partial_{tt} g\| + \Delta t \max_{t \in [0, T]} (\|\partial_t g\| + (v^2)^{1/2} \|\partial_{xt} \rho\|).
\end{aligned}$$

□

Subtracting the numerical scheme (2.10) from (3.48) and using the linearity, one gets the following error equations,

$$\left(\frac{\xi_\rho^{n+1} - \xi_\rho^n}{\Delta t}, \phi \right) + a_h(\xi_g^n, \phi) = \left(\frac{\eta_\rho^{n+1} - \eta_\rho^n}{\Delta t}, \phi \right) + a_h(\eta_g^n, \phi) + (\tau_\rho^n, \phi), \quad (3.51a)$$

$$\begin{aligned}
\left(\frac{\xi_g^{n+1} - \xi_g^n}{\Delta t}, \psi \right) + \frac{1}{\varepsilon} b_{h,v}(\xi_g^n, \psi) - \frac{v}{\varepsilon^2} d_h(\xi_\rho^{n+1}, \psi) + \frac{1}{\varepsilon^2} (\xi_g^{n+1}, \psi) & \quad (3.51b) \\
= \left(\frac{\eta_g^{n+1} - \eta_g^n}{\Delta t}, \psi \right) + \frac{1}{\varepsilon} b_{h,v}(\eta_g^n, \psi) - \frac{v}{\varepsilon^2} d_h(\eta_\rho^{n+1}, \psi) + \frac{1}{\varepsilon^2} (\eta_g^{n+1}, \psi) + (\tau_g^n, \psi)
\end{aligned}$$

for any test functions $\phi(\cdot), \psi(\cdot, v) \in U_h^k$. We now take $\phi = \xi_\rho^{n+1}$ in (3.51a). In addition, we take $\psi = \varepsilon^2 \xi_g^{n+1}$ in (3.51b), integrate in v , and shift the index n to $n - 1$. The resulting two equations are summed up and give LHS = RHS, where

$$\text{LHS} = \Xi_\varepsilon(\xi_\rho^{n+1}, \xi_\rho^n, \xi_g^n, \xi_g^{n-1}) + \|\xi_g^n\|^2, \quad (3.52)$$

$$\text{RHS} = \left(\frac{\eta_\rho^{n+1} - \eta_\rho^n}{\Delta t}, \xi_\rho^{n+1} \right) + a_h(\eta_g^n, \xi_\rho^{n+1}) + (\tau_\rho^n, \xi_\rho^{n+1}) \quad (3.53)$$

$$+ \varepsilon^2 \left\langle \left(\frac{\eta_g^n - \eta_g^{n-1}}{\Delta t}, \xi_g^n \right) \right\rangle + \varepsilon \langle b_{h,v}(\eta_g^{n-1}, \xi_g^n) \rangle - \langle v d_h(\eta_\rho^n, \xi_g^n) \rangle + \langle (\eta_g^n, \xi_g^n) \rangle + \varepsilon^2 \langle (\tau_g^{n-1}, \xi_g^n) \rangle.$$

Now we apply Lemma 3.2 with $\phi = \xi_\rho$, $\psi = \xi_g$, and $m = n - 1$, and get

$$\begin{aligned}
\text{LHS} &\geq \frac{1}{2\Delta t} (\mathcal{E}_{n+1} - \mathcal{E}_n) \quad (3.54) \\
&\quad + \begin{cases} (\varepsilon - \alpha_2 \frac{\Delta t}{h}) \langle \frac{|v|}{2} \sum_i [\xi_g^n]_{i-\frac{1}{2}}^2 \rangle + (1 - \alpha_1 \frac{\Delta t}{h^2}) \|\xi_g^n\|^2, & \text{for } k \geq 1, \\ (\varepsilon - \frac{\alpha_2}{2} \frac{\Delta t}{h}) \langle \frac{|v|}{2} \sum_i [\xi_g^n]_{i-\frac{1}{2}}^2 \rangle + \|\xi_g^n\|^2, & \text{for } k = 0. \end{cases}
\end{aligned}$$

Next we want to estimate RHS. This will be proceeded for the proposed scheme with the alternating and the central flux (2.9), respectively.

3.2.2 To estimate RHS (3.53) with the alternating flux

Without loss of generality, we consider using the alternating flux $\widehat{\langle vg \rangle} = \langle vg \rangle^-$, $\hat{\rho} = \rho^+$ in the proposed scheme (2.10). For this case, we take $\Pi_\rho = \pi_h^+$ and $\Pi_g = \pi_h^-$ in error estimates. With such choices and $\partial_x \xi_\rho^{n+1}, \partial_x (v \xi_g^n)(\cdot, v) \in U_h^{k-1}$, one has $\int_{I_i} \langle v \eta_g^n \rangle \partial_x \xi_\rho^{n+1} dx = \int_{I_i} \eta_\rho^n \partial_x \langle v \xi_g^n \rangle dx = 0$, and $\langle v \eta_g^n \rangle_{i-\frac{1}{2}}^- = (\eta_\rho^n)_{i-\frac{1}{2}}^+ = 0, \forall i$, therefore

$$a_h(\eta_g^n, \xi_\rho^{n+1}) = \langle v d_h(\eta_\rho^n, \xi_g^n) \rangle = 0. \quad (3.55)$$

We now turn to the term containing $b_{h,v}$ in (3.53). With $\langle \xi_g^n \rangle = 0, \forall n$ in (3.45), and the choice of Π_g , we have

$$\begin{aligned} \langle b_{h,v}(\eta_g^{n-1}, \xi_g^n) \rangle &= \langle (\mathcal{D}_{h,v}(\eta_g^{n-1}; v), \xi_g^n) \rangle \\ &= -\langle \sum_i \left(\int_{I_i} v \eta_g^{n-1} \partial_x \xi_g^n dx \right) + \sum_i (\widetilde{v \eta_g^{n-1}})_{i-\frac{1}{2}} [\xi_g^n]_{i-\frac{1}{2}} \rangle = -\langle \sum_i (\widetilde{v \eta_g^{n-1}})_{i-\frac{1}{2}} [\xi_g^n]_{i-\frac{1}{2}} \rangle \\ &= -\langle \sum_i (v \eta_g^{n-1})_{i-\frac{1}{2}}^- [\xi_g^n]_{i-\frac{1}{2}} \rangle + -\langle \sum_i (v \eta_g^{n-1})_{i-\frac{1}{2}}^+ [\xi_g^n]_{i-\frac{1}{2}} \rangle = -\langle \sum_i (v \eta_g^{n-1})_{i-\frac{1}{2}}^+ [\xi_g^n]_{i-\frac{1}{2}} \rangle. \end{aligned}$$

This can be further estimated as below, with any $\sigma \in (0, 1)$

$$\begin{aligned} |\langle b_{h,v}(\eta_g^{n-1}, \xi_g^n) \rangle| &\leq (1-\sigma) \langle \frac{|v|}{2} \sum_i [\xi_g^n]_{i-\frac{1}{2}}^2 \rangle_- + \frac{1}{4(1-\sigma)} \langle 2|v| \sum_i ((\eta_g^{n-1})_{i-\frac{1}{2}}^+)^2 \rangle_- \\ &\leq (1-\sigma) \langle \frac{|v|}{2} \sum_i [\xi_g^n]_{i-\frac{1}{2}}^2 \rangle + \frac{C_*}{(1-\sigma)} h^{2k+1}. \end{aligned} \quad (3.56)$$

The estimate in (3.14) is used to get the last inequality. We further note that

$$\begin{aligned} \left\| \frac{\eta_\rho^{n+1} - \eta_\rho^n}{\Delta t} \right\|^2 &= \int_{\Omega_x} \left| \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} (I - \Pi_\rho) \partial_t \rho(x, s) ds \right|^2 dx \\ &\leq \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \|(I - \Pi_\rho) \partial_t \rho(\cdot, s)\|^2 ds \\ &\leq C \max_{t \in [0, T]} \|\partial_t \rho(\cdot, t)\|_{H^{k+1}(\Omega_x)}^2 h^{2k+2} = C_* h^{2k+2}, \end{aligned} \quad (3.57)$$

and similarly

$$\left\| \frac{\eta_g^n - \eta_g^{n-1}}{\Delta t} \right\|^2 \leq C \max_{t \in [0, T]} \|\partial_t g(\cdot, t)\|_{H^{k+1}(\Omega_x)}^2 h^{2k+2} = C_* h^{2k+2}. \quad (3.58)$$

With these and the estimates on the truncation errors, we have

$$\begin{aligned} &\left| \left(\frac{\eta_\rho^{n+1} - \eta_\rho^n}{\Delta t}, \xi_\rho^{n+1} \right) + (\tau_\rho^n, \xi_\rho^{n+1}) \right| \\ &\leq \|\xi_\rho^{n+1}\|^2 + \frac{1}{4} \left\| \frac{\eta_\rho^{n+1} - \eta_\rho^n}{\Delta t} + \tau_\rho^n \right\|^2 \leq \|\xi_\rho^{n+1}\|^2 + C_*(h^{2k+2} + \Delta t^2), \end{aligned} \quad (3.59)$$

and

$$\begin{aligned} &|\varepsilon^2 \langle \left(\frac{\eta_g^n - \eta_g^{n-1}}{\Delta t}, \xi_g^n \right) \rangle + \langle (\eta_g^n, \xi_g^n) \rangle + \varepsilon^2 \langle (\tau_g^{n-1}, \xi_g^n) \rangle| \\ &\leq \varepsilon^2 \|\xi_g^n\|^2 + \frac{1}{4} \varepsilon^2 \left\| \frac{\eta_g^n - \eta_g^{n-1}}{\Delta t} \right\|^2 + (1-\sigma) \|\xi_g^n\|^2 + \frac{1}{4(1-\sigma)} (\|\varepsilon^2 \tau_g^{n-1}\|^2 + \|\eta_g^n\|^2) \\ &\leq \varepsilon^2 \|\xi_g^n\|^2 + (1-\sigma) \|\xi_g^n\|^2 + \varepsilon^2 C_* h^{2k+2} + \frac{C_*}{(1-\sigma)} ((1+\varepsilon^4) \Delta t^2 + h^{2k+2}). \end{aligned} \quad (3.60)$$

Now we combine the estimates in (3.55), (3.56), (3.59), (3.60), and get

$$|\text{RHS}| \leq \mathcal{E}_{n+1} + (1-\sigma) \|\xi_g^n\|^2 + (1-\sigma) \varepsilon \langle \frac{|v|}{2} \sum_i [\xi_g^n]_{i-\frac{1}{2}}^2 \rangle + C_* \tau, \quad (3.61)$$

where

$$\tau = (1+\varepsilon^2) h^{2k+2} + \Delta t^2 + \frac{1}{1-\sigma} ((1+\varepsilon^4) \Delta t^2 + h^{2k+2} + \varepsilon h^{2k+1}). \quad (3.62)$$

3.2.3 To estimate RHS (3.53) with the central flux

Now we consider the proposed scheme (2.10) with the central flux $\widehat{\langle vg \rangle} = \{\langle vg \rangle\}$, $\hat{\rho} = \{\rho\}$. For this case, we take $\Pi_\rho = \pi_h$ and $\Pi_g = \pi_h$. With such choices, there is

$$\begin{aligned} \text{RHS} &= a_h(\eta_g^n, \xi_\rho^{n+1}) + (\tau_\rho^n, \xi_\rho^{n+1}) + \varepsilon \langle b_{h,v}(\eta_g^{n-1}, \xi_g^n) \rangle - \langle v d_h(\eta_\rho^n, \xi_g^n) \rangle + \varepsilon^2 \langle (\tau_g^{n-1}, \xi_g^n) \rangle \\ &= (\tau_\rho^n, \xi_\rho^{n+1}) + \varepsilon^2 \langle (\tau_g^{n-1}, \xi_g^n) \rangle - \sum_i \left(\langle \{v \eta_g^n\} [\xi_\rho^{n+1}] \rangle + \varepsilon \langle \widetilde{\langle v \eta_g^{n-1} \rangle} [\xi_g^n] \rangle + \langle v \{ \eta_\rho^n \} [\xi_g^n] \rangle \right)_{i-\frac{1}{2}}. \end{aligned}$$

Moreover, one can use the estimates in (3.14), the estimates for the truncation errors, the inverse inequality (3.12a), and have

$$\begin{aligned} |(\tau_\rho^n, \xi_\rho^{n+1}) + \varepsilon^2 \langle (\tau_g^{n-1}, \xi_g^n) \rangle| &\leq \frac{1}{2} \|\xi_\rho^{n+1}\|^2 + \frac{1-\sigma}{2} \|\xi_g^n\|^2 + C_* \Delta t^2 + \frac{C_*}{1-\sigma} (1 + \varepsilon^4) \Delta t^2, \\ \left| \sum_i \langle \{v \eta_g^n\} [\xi_\rho^{n+1}] \rangle_{i-\frac{1}{2}} \right| &\leq \frac{h}{4C_{\text{inv}}} \sum_i [\xi_\rho^{n+1}]_{i-\frac{1}{2}}^2 + \frac{C_{\text{inv}}}{h} \sum_i \{v \eta_g^n\}_{i-\frac{1}{2}}^2 \leq \frac{1}{2} \|\xi_\rho^{n+1}\|^2 + C_* h^{2k}, \\ \left| \varepsilon \sum_i \langle \widetilde{\langle v \eta_g^{n-1} \rangle} [\xi_g^n] \rangle_{i-\frac{1}{2}} \right| &\leq (1-\sigma) \varepsilon \left\langle \frac{|v|}{2} \sum_i [\xi_g^n]_{i-\frac{1}{2}}^2 \right\rangle + \frac{C_*}{1-\sigma} \varepsilon h^{2k+1}, \\ \sum_i \langle v \{ \eta_\rho^n \} [\xi_g^n] \rangle_{i-\frac{1}{2}} &\leq \frac{1-\sigma}{2} \|\xi_g^n\|^2 + \frac{C_*}{1-\sigma} h^{2k}, \end{aligned}$$

therefore RHS can be estimated as (3.61) where

$$\tau = h^{2k} + \Delta t^2 + \frac{1}{1-\sigma} ((1 + \varepsilon^4) \Delta t^2 + h^{2k} + \varepsilon h^{2k+1}). \quad (3.63)$$

3.2.4 To estimate \mathcal{E}_1

With the discrete energy defined in the analysis, we also need to estimate \mathcal{E}_1 . To achieve this, we start with (3.51a), $n = 0$, and have

$$\begin{aligned} \|\xi_\rho^1\| &= \sup_{0 \neq \phi \in U_h^k} \frac{(\xi_\rho^1, \phi)}{\|\phi\|} = \sup_{0 \neq \phi \in U_h^k} \frac{(\eta_\rho^1 + e_\rho^0, \phi) - \Delta t a_h(e_g^0, \phi) + \Delta t (\tau_\rho^0, \phi)}{\|\phi\|} \\ &\leq \|\eta_g^1\| + \|e_\rho^0\| + \Delta t \left(\|\tau_\rho^0\| + \sup_{0 \neq \phi \in U_h^k} \frac{a_h(e_g^0, \phi)}{\|\phi\|} \right). \end{aligned}$$

The first two terms can be estimated based on the property of the projections in (3.14), that is, $\|\eta_\rho^1\| = \|\rho^1 - \Pi_\rho \rho^1\| = C_* h^{k+1}$ and $\|e_\rho^0\| = \|\rho^0 - \pi_h \rho^0\| = C_* h^{k+1}$. From Lemma 3.8, the truncation error τ_ρ^0 can be controlled by $\|\tau_\rho^0\| \leq C_* \Delta t$. To estimate the last term, we use $e_g^0 = g^0 - \pi_h g^0$ and (3.14), and have

$$|a_h(e_g^0, \phi)| = \left| \sum_i \widehat{\langle v e_g^0 \rangle}_{i-\frac{1}{2}} [\phi]_{i-\frac{1}{2}} \right| \leq \left(h^{-1} \sum_i \widehat{\langle v e_g^0 \rangle}_{i-\frac{1}{2}}^2 \right)^{1/2} \left(h \sum_i [\phi]_{i-\frac{1}{2}}^2 \right)^{1/2} \leq C_* h^k \|\phi\|,$$

for any $\phi \in U_h^k$. Note that $\Delta t \leq \sigma \Delta t_{stab}$ implies $\Delta t = C(h^2 + \varepsilon h)$, with this, we finally have

$$\begin{aligned} \mathcal{E}_1 &= \|\xi_\rho^1\|^2 + \varepsilon^2 \|\xi_g^0\|^2 = \|\xi_\rho^1\|^2 + \varepsilon^2 \|\Pi_g g^0 - g^0 + g^0 - \pi_h g^0\|^2 \\ &\leq C_* (\Delta t^2 h^{2k} + (1 + \varepsilon^2) h^{2k+2} + \Delta t^4) \leq C_* ((1 + \varepsilon^2) h^{2k+2} + \Delta t^4). \end{aligned}$$

3.2.5 The final step for the error estimates

We now combine the bounds of LHS and RHS in (3.54), (3.61) with (3.62) for the alternating flux, or with (3.63) for the central flux. Together with (3.35), we have

$$\frac{1}{2\Delta t} (\mathcal{E}_{n+1} - \mathcal{E}_n) \leq \mathcal{E}_{n+1} + \hat{\gamma} \|\xi_g^n\|^2 + C_* \tau \quad (3.64)$$

with

$$\hat{\gamma} = \begin{cases} \max\left(-\sigma\varepsilon + \alpha_2 \frac{\Delta t}{h}, 0\right) \frac{\alpha_3}{h} + \left(-\sigma + \alpha_1 \frac{\Delta t}{h^2}\right) & \text{for } k \geq 1 \\ \max\left(-\sigma\varepsilon + \frac{\alpha_2}{2} \frac{\Delta t}{h}, 0\right) \frac{\alpha_3}{h} - \sigma & \text{for } k = 0. \end{cases} \quad (3.65)$$

As long as $\hat{\gamma} \leq 0$, that is,

$$\Delta t \leq \sigma \Delta t_{stab}, \quad (3.66)$$

where Δt_{stab} comes from the time step constraint in (3.31) for numerical stability, then

$$(1 - 2\Delta t)\mathcal{E}_{n+1} \leq \mathcal{E}_n + (2\Delta t)C_*\tau. \quad (3.67)$$

Assume $\Delta t < \frac{1}{2}$. Define $\Theta_n = \mathcal{E}_n(1 - 2\Delta t)^n$, then $\Theta_{n+1} \leq \Theta_n + 2\Delta t(1 - 2\Delta t)^n C_*\tau$. With mathematical induction, and the estimate of \mathcal{E}_1 in Section 3.2.4, one gets

$$\Theta_n \leq 2\Delta t C_*\tau((1 - 2\Delta t) + \dots + (1 - 2\Delta t)^{n-1}) + \Theta_1 \leq (C_*\tau + \mathcal{E}_1)(1 - 2\Delta t) \leq C_*\tau, \quad (3.68)$$

that is $\mathcal{E}_n(1 - 2\Delta t)^n \leq C_*\tau$. Moreover, with $n : \Delta t n \leq T$, there is

$$\mathcal{E}_n \leq (1 - 2\Delta t)^{-n} C_*\tau \leq e^{2\Delta t n} C_*\tau \leq C_* e^{2T} \tau = C_*\tau. \quad (3.69)$$

On the other hand, the properties of the projection operators (3.14) ensure $\|\eta_\rho^n\| = C_* h^{k+1}$ and $\|\eta_g^{n-1}\| = C_* h^{k+1}$, therefore

$$\|e_\rho^n\|^2 + \varepsilon^2 \|e_g^{n-1}\|^2 \leq 2(\|\eta_\rho^n\|^2 + \varepsilon^2 \|\eta_g^{n-1}\|^2 + \mathcal{E}_n) \leq C_*(\tau + (1 + \varepsilon^2)h^{2k+2}). \quad (3.70)$$

We now can conclude the main error estimates in Theorem 3.6 by further utilizing the forms of τ in (3.62) and (3.63).

3.3 Rigorous asymptotic analysis

In [10], a formal asymptotic analysis was performed for the proposed methods, showing that when $\varepsilon \rightarrow 0$, the limiting schemes are consistent discretizations for the limiting heat equation. In this section, we want to establish this asymptotic preserving property rigorously for DG-IMEX1 using tools from functional analysis.

To explicitly indicate how the numerical solutions depend on the mesh parameter Δt , h , and on ε , we use $\rho_{\varepsilon, \Delta t, h}^n$ and $g_{\varepsilon, \Delta t, h}^n$ to denote ρ_h^n and g_h^n in this section, unless otherwise specified. For the initial condition $\rho_\varepsilon(x, t = 0)$ and $g_\varepsilon(x, v, t = 0)$, two assumptions are made which are mild and reasonable.

(A1) At $t = 0$,

$$\rho_\varepsilon \rightharpoonup \rho_0, \quad \langle v g_\varepsilon \rangle \rightharpoonup \langle v g_0 \rangle \text{ in } L^2(\Omega_x), \quad \text{as } \varepsilon \rightarrow 0. \quad (3.71)$$

Here “ \rightharpoonup ” stands for weak convergence.

(A2)

$$\sup_\varepsilon (\|\rho_\varepsilon\|)_{t=0} < \infty, \quad \sup_\varepsilon (\|g_\varepsilon\|)_{t=0} < \infty. \quad (3.72)$$

In the following analysis, the index k for the discrete space U_h^k , hence in the numerical method, is fixed. Let $\{\Psi_j\}_{j=1}^{N_k}$ denote an orthonormal basis of U_h^k with respect to the standard L^2 inner product. We define

$$\rho_{\Delta t, h}^0 = \pi_h \rho_0|_{t=0}, \quad q_{\Delta t, h}^0 = \pi_h \langle v g_0 \rangle|_{t=0}, \quad (3.73)$$

and also denote $q_{\varepsilon, \Delta t, h}^n = \langle v g_{\varepsilon, \Delta t, h}^n \rangle$, $\forall n$. Below we will discuss some properties of the numerical solution at $t = t^n$ with $n = 0, 1$, before turning to the main result in Theorem 3.10.

Lemma 3.9. Under the assumptions (A1) and (A2), we have

- (i) $\lim_{\varepsilon \rightarrow 0} \rho_{\varepsilon, \Delta t, h}^0 = \rho_{\Delta t, h}^0$, and $\lim_{\varepsilon \rightarrow 0} q_{\varepsilon, \Delta t, h}^0 = q_{\Delta t, h}^0$. The convergence is in any norm.
- (ii) $\sup_\varepsilon \|\rho_{\varepsilon, \Delta t, h}^1\| < \infty$, and $\sup_\varepsilon \|q_{\varepsilon, \Delta t, h}^0\| < \infty$.

Proof. First note that U_h^k is finite dimensional, then a sequence in U_h^k , if converges, will converge in any norm. To prove (i), based on assumption (A1), we have $\lim_{\varepsilon \rightarrow \infty} (\rho_\varepsilon, \psi) = (\rho_0, \psi)$, $\forall \psi \in L^2(\Omega_x)$ at $t = 0$, therefore as $\varepsilon \rightarrow 0$,

$$\rho_{\varepsilon, \Delta t, h}^0 = \pi_h \rho_\varepsilon|_{t=0} = \sum_{j=1}^{N_k} (\rho_\varepsilon|_{t=0}, \Psi_j) \Psi_j \rightarrow \sum_{j=1}^{N_k} (\rho_0|_{t=0}, \Psi_j) \Psi_j = \pi_h \rho_0|_{t=0} = \rho_{\Delta t, h}^0.$$

Similarly, the second half of (i) can be proved.

To prove (ii), recall that $g_{\varepsilon, \Delta t, h}^0 = \pi_h g_\varepsilon|_{t=0}$, then

$$\|g_{\varepsilon, \Delta t, h}^0\|^2 = \langle \|g_{\varepsilon, \Delta t, h}^0\|^2 \rangle = \langle \sum_{j=1}^{N_k} (g_\varepsilon|_{t=0}, \Psi_j)^2 \rangle \leq (\|g_\varepsilon\|^2)|_{t=0} \sum_{j=1}^{N_k} \|\Psi_j\|^2.$$

Now with assumption (A2), we have $\sup_\varepsilon \|g_{\varepsilon, \Delta t, h}^0\| < \infty$. Similarly, one can show that

$$\sup_\varepsilon \|\rho_{\varepsilon, \Delta t, h}^0\| \leq (\|\rho_\varepsilon\|)|_{t=0} \left(\sum_{j=1}^{N_k} \|\Psi_j\|^2 \right)^{1/2} < \infty. \quad (3.74)$$

What remains is to establish the boundedness of $\sup_\varepsilon \|\rho_{\varepsilon, \Delta t, h}^1\|$.

From (2.10a) with $n = 0$,

$$\|\rho_{\varepsilon, \Delta t, h}^1\| \leq \|\rho_{\varepsilon, \Delta t, h}^0\| + \Delta t \sup_{0 \neq \phi \in U_h^k} \frac{a_h(g_{\varepsilon, \Delta t, h}^0, \phi)}{\|\phi\|}. \quad (3.75)$$

Without loss of generality, in $a_h(\cdot, \cdot)$ defined in (2.6a), we consider $\widehat{\langle vg \rangle} = \langle vg \rangle^-$. First we have

$$\begin{aligned} \|\langle vg_{\varepsilon, \Delta t, h}^0 \rangle\|^2 &= \left\| \sum_{j=1}^{N_k} (\langle vg_\varepsilon \rangle|_{t=0}, \Psi_j) \Psi_j \right\|^2 = \sum_{j=1}^{N_k} (\langle vg_\varepsilon \rangle|_{t=0}, \Psi_j)^2 \\ &\leq (\|\langle vg_\varepsilon \rangle\|^2)|_{t=0} \sum_{j=1}^{N_k} \|\Psi_j\|^2 \leq \langle v^2 \rangle (\|g_\varepsilon\|^2)|_{t=0} \sum_{j=1}^{N_k} \|\Psi_j\|^2, \end{aligned} \quad (3.76)$$

and

$$\begin{aligned} |\langle vg_{\varepsilon, \Delta t, h}^0 \rangle_{i-\frac{1}{2}}^-| &= \left| \sum_{j=1}^{N_k} (\langle vg_\varepsilon \rangle|_{t=0}, \Psi_j) \Psi_j(x_{i-\frac{1}{2}}^-) \right| \\ &\leq \langle v^2 \rangle^{1/2} (\|g_\varepsilon\|)|_{t=0} \sum_{j=1}^{N_k} \|\Psi_j\| |\Psi(x_{i-\frac{1}{2}}^-)|. \end{aligned} \quad (3.77)$$

Based on the definition of $a_h(\cdot, \cdot)$, and inverse inequalities in (3.12), we further get

$$\begin{aligned} a_h(g_{\varepsilon, \Delta t, h}^0, \phi) &= -(\langle vg_{\varepsilon, \Delta t, h}^0 \rangle, \partial_x \phi) - \sum_i \langle \widehat{vg_{\varepsilon, \Delta t, h}^0} \rangle_{i-\frac{1}{2}} [\phi]_{i-\frac{1}{2}} \\ &\leq \|\langle vg_{\varepsilon, \Delta t, h}^0 \rangle\| \|\partial_x \phi\| + \sum_i |\langle \widehat{vg_{\varepsilon, \Delta t, h}^0} \rangle_{i-\frac{1}{2}}| |\phi]_{i-\frac{1}{2}}| \\ &\leq \frac{C(k)}{h} \|\phi\| \left(\|\langle vg_{\varepsilon, \Delta t, h}^0 \rangle\| + \left(h \sum_i (\langle \widehat{vg_{\varepsilon, \Delta t, h}^0} \rangle_{i-\frac{1}{2}}^-)^2 \right)^{1/2} \right). \end{aligned} \quad (3.78)$$

Now we can combine (3.74)-(3.78) as well as the boundedness assumption (A2), and conclude

$$\sup_\varepsilon \|\rho_{\varepsilon, \Delta t, h}^1\| \leq C(k, \Delta t, h, \langle v^2 \rangle) \left(\sup_\varepsilon \|\rho_\varepsilon\| + \sup_\varepsilon \|g_\varepsilon\| \right) |_{t=0} < \infty.$$

□

Theorem 3.10. *Let c_0 be any fixed positive constant, $c_0 \in (0, 1)$. Under the assumptions (A1) and (A2) on the initial data, and under the following condition on the time step Δt ,*

$$\Delta t < \Delta t_{stab, c_0} = \frac{h}{\alpha_1 + \alpha_2 \alpha_3} \left((1 - c_0)h + \min(\varepsilon, \frac{(1 - c_0)\alpha_2 h}{\alpha_1})\alpha_3 \right), \quad (3.79)$$

with $\alpha_i, i = 1, 2, 3$ defined in (3.13), we have

$$\lim_{\varepsilon \rightarrow 0} \rho_{\varepsilon, \Delta t, h}^n = \rho_{\Delta t, h}^n, \quad q_{\varepsilon, \Delta t, h}^n \rightharpoonup q_{\Delta t, h}^n \text{ in } L^2(\Omega_x) \text{ as } \varepsilon \rightarrow 0, \quad \forall n \geq 1 \quad (3.80)$$

for some $\rho_{\Delta t, h}^n, q_{\Delta t, h}^n \in U_h^k$, and the limits satisfy

$$\left(\frac{\rho_{\Delta t, h}^{n+1} - \rho_{\Delta t, h}^n}{\Delta t}, \phi \right) - \sum_i \int_{I_i} q_{\Delta t, h}^n \partial_x \phi dx - \sum_i (\widehat{q_{\Delta t, h}^n})_{i-\frac{1}{2}} [\phi]_{i-\frac{1}{2}} = 0, \quad \forall \phi \in U_h^k, \quad (3.81a)$$

$$(q_{\Delta t, h}^{n+1}, \psi) = \langle v^2 \rangle \left(\sum_i \int_{I_i} \rho_{\Delta t, h}^{n+1} \partial_x \psi dx + \sum_i (\widehat{\rho_{\Delta t, h}^{n+1}})_{i-\frac{1}{2}} [\psi]_{i-\frac{1}{2}} \right), \quad \forall \psi \in U_h^k, \quad (3.81b)$$

for $n \geq 0$, with the initial data given by (3.73).

The limiting scheme (3.81), though being implicitly defined, is intrinsically explicit if one first solves $\rho_{\Delta t, h}^{n+1}$ then $g_{\Delta t, h}^{n+1}$ in actual implementation. Note that this limiting scheme with any fixed k is a consistent scheme for the limiting heat equation $\partial_t \rho = \partial_x (\langle v^2 \rangle \partial_x \rho)$ in its first order form

$$\partial_t \rho + \partial_x q = 0, \quad q = -\langle v^2 \rangle \partial_x \rho. \quad (3.82)$$

In fact, the spatial discretization in the limiting scheme is exactly the local DG spatial discretization for the heat equation studied in [8].

Proof. We start with revisiting γ in (3.36) and (3.37) from the stability analysis. By requiring $\gamma \leq -c_0$, we obtain the condition (3.79) on the time step Δt . It is easy to see $\Delta t_{stab, c_0} < \Delta t_{stab}$ (In fact, $\lim_{c_0 \rightarrow 0} \Delta t_{stab, c_0} = \Delta t_{stab}$.) The remaining of the proof consists of two steps.

Step 1: In this step, we want to show

$$\sup_{\varepsilon} \|\rho_{\varepsilon, \Delta t, h}^n\| < \infty, \quad \sup_{\varepsilon} \|g_{\varepsilon, \Delta t, h}^n\| < \infty, \quad \forall n \geq 1, \quad (3.83)$$

when Δt and h satisfy (3.79). With such mesh parameters, equation (3.36) turns to

$$\frac{1}{2\Delta t} \left((\|\rho_{\varepsilon, \Delta t, h}^{n+1}\|^2 + \varepsilon^2 \|g_{\varepsilon, \Delta t, h}^n\|^2) - (\|\rho_{\varepsilon, \Delta t, h}^n\|^2 + \varepsilon^2 \|g_{\varepsilon, \Delta t, h}^{n-1}\|^2) \right) \leq -c_0 \|g_{\varepsilon, \Delta t, h}^n\|^2, \quad (3.84)$$

that is

$$\|\rho_{\varepsilon, \Delta t, h}^{n+1}\|^2 + (2\Delta t c_0 + \varepsilon^2) \|g_{\varepsilon, \Delta t, h}^n\|^2 \leq \|\rho_{\varepsilon, \Delta t, h}^n\|^2 + \varepsilon^2 \|g_{\varepsilon, \Delta t, h}^{n-1}\|^2. \quad (3.85)$$

On the other hand, $\Delta t_{stab, c_0} < \Delta t_{stab}$, and this implies the stability estimate (3.30). Combining (3.85), (3.30), the boundedness of $\sup_{\varepsilon} \|\rho_{\varepsilon, \Delta t, h}^1\|, \sup_{\varepsilon} \|g_{\varepsilon, \Delta t, h}^0\|$ in Lemma 3.9, we will obtain (3.83).

Step 2: Now we would like to establish the asymptotic behavior in (3.80), as well as the fact that the limits $\rho_{\Delta t, h}^n$ and $q_{\Delta t, h}^n$ satisfy (3.81) with the initial data (3.73).

First of all, it is easy to see that to obtain (3.80), it is equivalent to show

$$\lim_{m \rightarrow \infty} \rho_{\varepsilon_m, \Delta t, h}^n = \rho_{\Delta t, h}^n, \quad q_{\varepsilon_m, \Delta t, h}^n \rightharpoonup q_{\Delta t, h}^n \text{ in } L^2(\Omega_x) \text{ as } m \rightarrow \infty, \quad \forall n \geq 1 \quad (3.86)$$

where $\{\varepsilon_m\}_{m=1}^{\infty}$ is any sequence such that $\lim_{m \rightarrow \infty} \varepsilon_m = 0$. Given that U_h^k is finite dimensional, the boundedness of $\sup_m \|\rho_{\varepsilon_m, \Delta t, h}^n\|$ from (3.83) implies that there is a subsequence $\{\rho_{\varepsilon_{m_r}, \Delta t, h}^n\}_{r=1}^{\infty}$ converging in U_h^k under any norm as $r \rightarrow \infty$. Let's denote the limit as $\rho_{\Delta t, h}^n \in U_h^k$.

Now we turn to $\{q_{\varepsilon_m, \Delta t, h}^n\}_{m=1}^{\infty}$. For the simplicity of notations, this sequence will be denoted as $\{q_{\varepsilon_m}^n\}_{m=1}^{\infty} = \{\langle v g_{\varepsilon_m}^n \rangle\}_{m=1}^{\infty}$ in the present paragraph. For each function g_{ε_m} , it can be written as $g_{\varepsilon_m}^n(x, v) = \sum_{j=1}^{N_k} \alpha_{\varepsilon_m}^{(j)}(v) \Psi_j(x)$.

In addition, we have $\|g_{\varepsilon_m}^n\| = \left(\sum_{j=1}^{N_k} \|\alpha_{\varepsilon_m}^{(j)}\|_{L^2(\Omega_v)}^2\right)^{1/2}$. This, in addition to the boundedness of $\sup_m \|g_{\varepsilon_m}^n\|$ in (3.83), indicates that $\sup_m \|\alpha_{\varepsilon_m}^{(j)}\|_{L^2(\Omega_v)}^2$, therefore $\sup_r \|\alpha_{\varepsilon_{m_r}}^{(j)}\|_{L^2(\Omega_v)}^2$ is bounded for any $j = 1, \dots, N_k$. As a Hilbert space, $L^2(\Omega_v)$ is weakly sequentially compact, that is, $\{\alpha_{\varepsilon_{m_r}}^{(j)}\}_{r=1}^\infty$ has a subsequence which is weakly convergent in $L^2(\Omega_v)$. Without loss of generality, this subsequence is still denoted as $\{\alpha_{\varepsilon_{m_r}}^{(j)}\}_{r=1}^\infty$, and the weak limit when $r \rightarrow \infty$ is denoted as $\alpha_0^{(j)} \in L^2(\Omega_v)$, $\forall j$. We now define $g_{\Delta t, h}^n(x, v) = \sum_{j=1}^{N_k} \alpha_0^{(j)}(v) \Psi_j(x)$, and $q_{\Delta t, h}^n = \langle v g_{\Delta t, h}^n \rangle = \sum_{j=1}^{N_k} \langle v \alpha_0^{(j)} \rangle \Psi_j(x)$. For any $\psi \in U_h^k$,

$$\lim_{r \rightarrow \infty} (q_{\varepsilon_{m_r}, \Delta t, h}^n, \psi) = \sum_{j=1}^{N_k} \left(\lim_{r \rightarrow \infty} \langle v \alpha_{\varepsilon_{m_r}}^{(j)} \rangle \right) (\Psi_j(x), \psi) = \sum_{j=1}^{N_k} \left(\langle v \alpha_0^{(j)} \rangle \right) (\Psi_j(x), \psi) = (q_{\Delta t, h}^n, \psi).$$

Up to now, we have shown that (3.80) holds for a subsequence of $\{\rho_{\varepsilon_{m_r}, \Delta t, h}^n\}_{r=1}^\infty$ and $\{q_{\varepsilon_{m_r}, \Delta t, h}^n\}_{r=1}^\infty$ as $r \rightarrow \infty$. Moreover, it is straightforward to see that the limits $\rho_{\Delta t, h}^n$ and $g_{\Delta t, h}^n$, $n = 1, 2, \dots$, satisfy (3.81) with the initial data (3.73). On the other hand, given the initial data (3.73), the solution to (3.73) at t^{n+1} , $n \geq 0$, is uniquely determined by first solving (3.81a) for $\rho_{\Delta t, h}^{n+1}$ and then solving (3.81b) for $q_{\Delta t, h}^{n+1}$. Finally, we can follow a standard contradiction argument and the uniqueness of the solution to (3.81) to conclude that (3.80) holds for the entire sequence. \square

Remark 3.11. The rigorous asymptotic analysis is established for the methods with the first order accuracy in time (2.10). When higher order temporal discretizations are used as in [10], one can follow the steps in this subsection to obtain a rigorous asymptotic analysis, as long as a stability estimate similar to (3.36) is available.

Acknowledgement. This project was initiated during the authors' participation at the ICERM Semester Program on "Kinetic Theory and Computation" in the fall of 2011. The authors want to thank for the generous support from the Institute. Part of the work was done at MFO in Oberwolfach during a Research in Pairs program. The first three authors appreciate the support and hospitality of the Institute.

References

- [1] C. BARDOS, F. GOLSE, AND D. LEVERMORE, *Fluid dynamic limits of kinetic equations. I. formal derivations*, Journal of Statistical Physics, 63 (1991), pp. 323–344.
- [2] C. BARDOS, R. SANTOS, AND R. SENTIS, *Diffusion approximation and computation of the critical size*, Transactions of the American Mathematical Society, 284 (1984), pp. 617–649.
- [3] S. BOSCARINO, L. PARESCHI, AND G. RUSSO, *Implicit-explicit Runge–Kutta schemes for hyperbolic systems and kinetic equations in the diffusion limit*, SIAM Journal on Scientific Computing, 35 (2013), pp. A22–A51.
- [4] S. BOSCARINO AND G. RUSSO, *Flux-explicit IMEX Runge–Kutta schemes for hyperbolic to parabolic relaxation problems*, SIAM Journal on Numerical Analysis, 51 (2013), pp. 163–190.
- [5] J. A. CARRILLO, T. GOUDON, P. LAFITTE, AND F. VECIL, *Numerical schemes of diffusion asymptotics and moment closures for kinetic equations*, Journal of Scientific Computing, 36 (2008), pp. 113–149.
- [6] C. CERCIGNANI, *The Boltzmann equation*, Springer, 1988.
- [7] P. G. CIARLET, *The finite element method for elliptic problems*, North Holland, 1975.
- [8] B. COCKBURN AND C.-W. SHU, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM Journal on Numerical Analysis, 35 (1998), pp. 2440–2463.
- [9] J.-L. GUERMOND AND G. KANSCHAT, *Asymptotic analysis of upwind discontinuous Galerkin approximation of the radiative transport equation in the diffusive limit*, SIAM Journal on Numerical Analysis, 48 (2010), pp. 53–78.
- [10] J. JANG, F. LI, J.-M. QIU, AND T. XIONG, *High order asymptotic preserving DG-IMEX schemes for discrete-velocity kinetic equations in a diffusive scaling*, <http://arxiv.org/abs/1306.0227>, submitted, (2013).

- [11] S. JIN, *Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review*, Lecture Notes for Summer School on Methods and Models of Kinetic Theory (M&MKT), Porto Ercole (Grosseto, Italy), (2010).
- [12] S. JIN AND C. LEVERMORE, *Numerical schemes for hyperbolic conservation laws with stiff relaxation terms*, Journal of Computational Physics, 126 (1996), pp. 449–467.
- [13] S. JIN, L. PARESCHI, AND G. TOSCANI, *Diffusive relaxation schemes for multiscale discrete-velocity kinetic equations*, SIAM Journal on Numerical Analysis, 35 (1998), pp. 2405–2439.
- [14] ———, *Uniformly accurate diffusive relaxation schemes for multiscale transport equations*, SIAM Journal on Numerical Analysis, 38 (2000), pp. 913–936.
- [15] A. KLAR, *An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit*, SIAM Journal on Numerical Analysis, 35 (1998), pp. 1073–1094.
- [16] P. LAFITTE AND G. SAMAËY, *Asymptotic-preserving projective integration schemes for kinetic equations in the diffusion limit*, SIAM Journal on Scientific Computing, 34 (2012), pp. 579–602.
- [17] E. W. LARSEN AND J. MOREL, *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes II*, Journal of Computational Physics, 83 (1989), pp. 212–236.
- [18] E. W. LARSEN, J. MOREL, AND W. F. MILLER JR, *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes*, Journal of Computational Physics, 69 (1987), pp. 283–324.
- [19] M. LEMOU AND L. MIEUSSENS, *A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit*, SIAM Journal on Scientific Computing, 31 (2010), pp. 334–368.
- [20] J. LIU AND L. MIEUSSENS, *Analysis of an asymptotic preserving scheme for linear kinetic equations in the diffusion limit*, SIAM Journal on Numerical Analysis, 48 (2010), pp. 1474–1491.
- [21] T.-P. LIU AND S.-H. YU, *Boltzmann equation: micro-macro decompositions and positivity of shock profiles*, Communications in Mathematical Physics, 246 (2004), pp. 133–179.
- [22] R. LOWRIE AND J. MOREL, *Methods for hyperbolic systems with stiff relaxation*, International Journal for Numerical Methods in Fluids, 40 (2002), pp. 413–423.
- [23] G. NALDI AND L. PARESCHI, *Numerical schemes for kinetic equations in diffusive regimes*, Applied Mathematics Letters, 11 (1998), pp. 29–35.
- [24] L. PARESCHI AND G. RUSSO, *Efficient asymptotic preserving deterministic methods for the Boltzmann equation*, AVT-194 RTO AVT/VKI, Models and Computational Methods for Rarefied Flows, Lecture Series held at the von Karman Institute, Rhode St. Genese, Belgium, (2011).
- [25] M. REYNA AND F. LI, *Operator bounds and time step conditions for DG and central DG methods*, Journal of Scientific Computing, in print, (2014).
- [26] L. SAINT-RAYMOND, *Hydrodynamic limits of the Boltzmann equation*, vol. 1971, Springer, 2009.
- [27] T. WARBURTON AND T. HAGSTROM, *Taming the CFL number for discontinuous Galerkin methods on structured meshes*, SIAM Journal on Numerical Analysis, 46 (2008), pp. 3151–3180.
- [28] Q. ZHANG AND C.-W. SHU, *Error estimates to smooth solutions of Runge–Kutta discontinuous Galerkin methods for scalar conservation laws*, SIAM Journal on Numerical Analysis, 42 (2004), pp. 641–666.
- [29] ———, *Stability analysis and a priori error estimates of the third order explicit Runge–Kutta discontinuous Galerkin method for scalar conservation laws*, SIAM Journal on Numerical Analysis, 48 (2010), pp. 1038–1063.