A High-Order Asymptotic-Preserving Scheme for Kinetic Equations Using Projective Integration — Source link ↗

Pauline Lafitte, Annelies Lejon, Giovanni Samaey

Institutions: CentraleSupélec

Related papers:

- A high-order asymptotic-preserving scheme for kinetic equations using projective integration

- Second-order accurate projective integrators for multiscale problems

- A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems

- Heterogeneous multiscale methods: A review

- Coarse projective kMC integration: forward/reverse initial and boundary value problems

A high-order asymptotic-preserving scheme for kinetic equations using projective integration

Pauline Lafitte
Annelies Lejon
Giovanni Samaey

A high-order asymptotic-preserving scheme for kinetic equations using projective integration

Pauline Lafitte[*]
Annelies Lejon[†]
Giovanni Samaey

Report TW 646, April 2014

Department of Computer Science, KU Leuven

Abstract

We investigate a high-order, fully explicit, asymptotic-preserving scheme for a kinetic equation with linear relaxation, both in the hydrodynamic and diffusive scalings in which a hyperbolic, resp. parabolic, limiting equation exists. The scheme first takes a few small (inner) steps with a simple, explicit method (such as direct forward Euler) to damp out the stiff components of the solution and estimate the time derivative of the slow components. These estimated time derivatives are then used in an (outer) Runge–Kutta method of arbitrary order. We show that, with an appropriate choice of inner step size, the time-step restriction on the outer time step is similar to the stability condition for the limiting macroscopic equation. Moreover, the number of inner time steps is also independent of the scaling parameter. We analyse stability and consistency, and illustrate with numerical results.

Keywords : kinetic equations, projective integration, higher order, asymptotic-preserving.

A high-order asymptotic-preserving scheme for kinetic equations using projective integration

April 24, 2014

Abstract

We investigate a high-order, fully explicit, asymptotic-preserving scheme for a kinetic equation with linear relaxation, both in the hydrodynamic and diffusive scalings in which a hyperbolic, resp. parabolic, limiting equation exists. The scheme first takes a few small (inner) steps with a simple, explicit method (such as direct forward Euler) to damp out the stiff components of the solution and estimate the time derivative of the slow components. These estimated time derivatives are then used in an (outer) Runge–Kutta method of arbitrary order. We show that, with an appropriate choice of inner step size, the time-step restriction on the outer time step is similar to the stability condition for the limiting macroscopic equation. Moreover, the number of inner time steps is also independent of the scaling parameter. We analyse stability and consistency, and illustrate with numerical results.

1 Introduction

In many applications (such as traffic flow, biology or physics), the system under study consists of a large number of interacting particles. One option is to simulate such systems at a microscopic level, via an agent-based description with great modelling detail. At a mesoscopic level, one can write a kinetic description that governs the evolution of the particle distribution in position-velocity space. Then, $f(x, v, t)$ represents the probability of finding a particle at position x, moving with velocity v at time t. Its evolution is governed by a kinetic equation,

$$\partial_t f^\varepsilon + v \partial_x f^\varepsilon = Q(f^\varepsilon), \tag{1}$$

in which the lefthand side describes free transport and $Q(f)$ embodies collisions (velocity changes). Equation (1) can be made dimensionless via a rescaling with respect to the characteristic length L, time T and velocity V scales

$$\tilde{x} = Lx \qquad \tilde{t} = Tt \qquad \tilde{v} = Vv.$$

The regimes in which we are interested are $L = VT\varepsilon^\gamma$, where ε is a positive constant and γ is an integer that indicates a hydrodynamic ($\gamma = 0$) or diffusive ($\gamma = 1$) scaling. (Details on the choice of scaling are in section 2.) This, omitting the tildes, results in the dimensionless equation

$$\partial_t f^\varepsilon + \frac{v}{\varepsilon^\gamma} \partial_x f^\varepsilon = \frac{Q(f^\varepsilon)}{\varepsilon^{\gamma+1}}. \tag{2}$$

In a diffusive or hydrodynamic scaling, one can usually obtain an approximate macroscopic partial differential equation (PDE) for a number of low-order moments of the particle distribution f (such as density, momentum, etc.). Improving upon the macroscopic approximation, however, is computationally expensive. Because of the stiffness of (2), explicit methods require an excessively small time-step for small values of ε, whereas implicit methods suffer from the high dimensionality of the problem.

There is currently a large research effort in the design of algorithms that are uniformly stable in ε and approach a scheme for the limiting equation when ε tends to 0; such schemes are called *asymptotic-preserving in the sense of Jin* [31]. We briefly review here some achievements, and refer to the cited references for more

details. In [28, 29, 35], separating the distribution f into its odd and even parts in the velocity variable results in a coupled system of transport equations where the stiffness appears only in the source term, allowing to use a time-splitting technique [49] with implicit treatment of the source term; see also related work in [31, 35, 34, 36]. Implicit-explicit (IMEX) schemes are an extensively studied technique to tackle this kind of problems [3, 17] (and references therein). Recent results in this setting were obtained by Dimarco *et al.* to deal with nonlinear collision kernels [13], and an extension to hyperbolic systems in a diffusive limit is given in [6]. A different point of view based on well-balanced methods was introduced by Gosse and Toscani [22, 23], see also [9, 8]. Discontinuous Galerkin schemes have also been developed [38, 1, 41, 42, 24], as well as regularization methods [27, 25]. When the collision operator allows for an explicit computation, an explicit scheme can be obtained subject to a classical diffusion CFL condition by splitting f into its mean value and the first-order fluctuations in a Chapman-Enskog expansion form [20]. Also, closure by moments [12, e.g.] can lead to reduced systems for which time-splitting provides new classes of schemes [10], see [44, 45, 40, 50] for more complete references on moment methods in general. Alternatively, a micro-macro decomposition based on a Chapman-Enskog expansion has been proposed [40], leading to a system of transport equations that allows to design a semi-implicit scheme without time splitting. An non-local procedure based on the quadrature of kernels obtained through pseudo-differential calculus was proposed in [4].

In [37], an alternative, fully explicit, asymptotic-preserving method was proposed, based on projective integration, which was introduced in [18] as an explicit method for stiff systems of ordinary differential equations (ODEs) that have a large gap between their fast and slow time scales; these methods fit within recent research efforts on numerical methods for multiscale simulation [14, 15, 32, 33]; see also [16, 48, 52] for related approaches. In projective integration, the fast modes, corresponding to the Jacobian eigenvalues with large negative real parts, decay quickly, whereas the slow modes correspond to eigenvalues of smaller magnitude and are the main contributions to the solution. Projective integration allows a stable yet explicit integration of such problems by first taking a few small (inner) steps with a simple, explicit method, until the transients corresponding to the fast modes have died out, and subsequently projecting (extrapolating) the solution forward in time over a large (outer) time step; a schematic representation of the scheme is given in figure 1. In [37], this method was shown to be asymptotic-preserving for kinetic equations in the diffusive
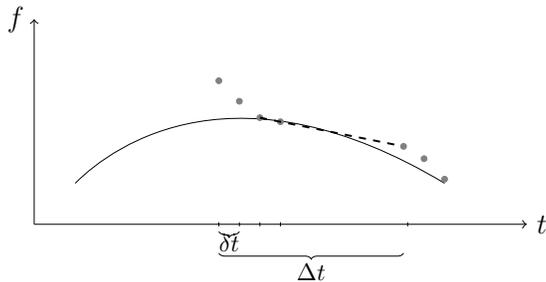

Figure 1: Illustration of first order projective integration.

scaling with a linear relaxation collision operator: given an adequate choice of the size of the inner time step, one can obtain a method that has a CFL-type step-size restriction on the outer time step, and requires a number of inner steps that is independent of ε. The computational cost of the method is thus independent of ε.

Many of the above-described methods have inherent limitations with respect to the order that can be achieved with the time discretisation, for instance due to the time-splitting. In this paper, we present a projective integration method that allows to attain arbitrary order accuracy in time. The generalisation is based on a modification of classical Runge–Kutta methods, and retains all advantages of the method in [37], i.e., it is fully explicit and asymptotic-preserving. Additionally, we significantly extend the analysis of the scheme. Specifically, the results in [37] are limited to the diffusive scaling, and to an equation that has a pure diffusion limiting equation. In this paper, we extend these results to model equations that result

2

in an advection-diffusion limit when ε tends to 0, and the analysis now covers both the diffusive and the hydrodynamic scaling. In [43], we discuss and illustrate how the method can be used in conjunction with a relaxation method [2, 30] to create a fully general, explicit time integration method for hyperbolic systems of conservation laws, also in multiple space dimensions. We remark that alternative approaches to obtain a higher-order projective integration scheme have been proposed in [39, 46].

The remainder of the paper is organized as follows. In section 2, we discuss the model problems that will be used during the numerical experiments. We then discuss the projective Runge-Kutta method (PRK) in section 3, and provide a result concerning its stability region. In section 4, we perform an analysis of the spectrum of the kinetic equations introduced in section 2, generalising the results obtained in [37] to the hydrodynamic scaling and to systems with macroscopic advection. The analysis also reveals how to choose the different method parameters of the projective Runge–Kutta method. In section 5 we give a consistency proof that shows the order of accuracy. We illustrate the results with some numerical experiments in section 6. Finally, section 7 contains a conclusion and outlook to future work.

2 Model problems

2.1 A simple kinetic equation

As a first model problem, we study a dimensionless scalar kinetic equation with linear relaxation in one space dimension,

$$\partial_t f^\varepsilon + \frac{v}{\varepsilon^\gamma}\partial_x f^\varepsilon = \frac{\mathcal{M}_v(u^\varepsilon) - f^\varepsilon}{\varepsilon^{\gamma+1}}, \tag{3}$$

modelling the evolution of a particle distribution function $f^\varepsilon(x, v, t)$ that gives the distribution of particles at a given position $x \in U = [-1, 1)$ with velocity $v \in V \subset \mathbb{R}$ at time $t > 0$, ε being a positive fixed constant. For the consistency analysis, we will impose periodic boundary conditions In the numerical experiments, we will also use Neumann boundary conditions. The parameter γ defines the scaling: when $\gamma = 0$, the scaling is called hydrodynamic; $\gamma = 1$ corresponds to a diffusive scaling. The righthand side represents a BGK collision operator [5] that models linear relaxation of f^ε towards a Maxwellian distribution $\mathcal{M}_v(u^\varepsilon)$, in which $u^\varepsilon(x, t) = \langle f^\varepsilon(x, v, t)\rangle$ is the density, obtained via averaging over the measured velocity space (V, μ), i.e.,

$$\langle \cdot \rangle = \int_V \cdot \, d\mu(v). \tag{4}$$

Let us now discuss the measured velocity space (V, μ) and the Maxwellian $\mathcal{M}_v$.

Velocity space   We consider odd-symmetric velocity spaces (V, μ) :

$$\begin{cases} \langle 1 \rangle = \int_V d\mu(v) = 1, \\ \langle h \rangle = \int_V h(v)d\mu(v) = 0 \text{ for any odd integrable function } h : V \longrightarrow \mathbb{R}, \\ \langle v^2 \rangle = \int_V v^2 d\mu(v) = d > 0. \end{cases}$$

We restrict ourselves to discretized velocity spaces of the form

$$V := \{v_j\}_{j=1}^J, \qquad d\mu(v) = \sum_{j=1}^J w_j \delta(v - v_j), \tag{5}$$

with J even, where the velocities satisfy $v_j = -v_{J-j}$ for all j, and w_j are appropriately chosen weights that satisfy $\sum_j^J w_j = 1$.

In the diffusive scaling ($\gamma = 1$), the discrete velocity space V results from applying a Gauss quadrature discretisation to (4) [11]. Throughout the analysis, we will consider a uniform symmetric discretisation of $V = (-1, 1)$, i.e., $v_j = (2j - J - 1)/J$, with $J/2 + 1 \leq j \leq J$; the weights are then defined as $w_j = 1/J$. In

our numerical experiments, we will use discretisations of (i) the velocity space $V = (-1, 1)$ endowed with the Lebesgue measure; and (ii) the velocity space $V = \mathbb{R}$ endowed with the Gaussian measure $d\mu(v) = (2\pi)^{-1/2} \exp(-v^2/2) dv$. Then, v_j are chosen as the roots of the Legendre, resp. Hermite, polynomial of degree $2J$, and the w_j are the corresponding quadrature weights. In the hyperbolic scaling ($\gamma = 0$), (V, μ) needs to satisfy the subcharacteristic condition (which ensures the positivity of the diffusion coefficient), see, e.g., [2, 43].

Maxwellian   Let us assume that the Maxwellian $\mathcal{M}_v$ satisfies (see, e.g., [2, 7])

$$\langle \mathcal{M}_v(u) \rangle = u, \qquad \langle v \mathcal{M}_v(u) \rangle = \varepsilon^\gamma A(u). \tag{6}$$

Throughout the analysis and numerical experiments, we will use

$$\mathcal{M}_v(u) = u + \varepsilon^\gamma \frac{A(u)}{v}. \tag{7}$$

In the analysis, we will restrict ourselves to the linear case, $A(u) = u$.

Let us now discuss the limiting macroscopic equation when ε tends to 0 by performing a Chapman-Enskog expansion,

$$f^\varepsilon = \mathcal{M}_v(u^\varepsilon) + \varepsilon g^\varepsilon, \tag{8}$$

with $\langle g^\varepsilon \rangle = 0$. Substituting (8) into the model equation (3) yields

$$\partial_t \left(\mathcal{M}_v(u^\varepsilon) + \varepsilon g^\varepsilon \right) + \frac{v}{\varepsilon^\gamma} \partial_x \left(\mathcal{M}_v(u^\varepsilon) + \varepsilon g^\varepsilon \right) = -\frac{g^\varepsilon}{\varepsilon^\gamma}. \tag{9}$$

Then, taking the mean over velocity space and using (6), we obtain

$$\partial_t u^\varepsilon + \partial_x (A(u^\varepsilon)) + \varepsilon^{1-\gamma} \langle v \partial_x g^\varepsilon \rangle = 0.$$

The last term on the lefthand side can be approximated by considering the terms in (9) of order $O(1/\varepsilon^\gamma)$, from which we obtain $g^\varepsilon = -v \partial_x \mathcal{M}_v(u^\varepsilon) + O(\varepsilon)$. This gives rise to

$$\partial_t u^\varepsilon + \partial_x (A(u^\varepsilon)) = \varepsilon^{1-\gamma} d \, \partial_{xx} u^\varepsilon + O(\varepsilon^2), \qquad d = \langle v^2 \rangle \tag{10}$$

Depending on the scaling, we thus obtain a hyperbolic advection equation ($\gamma = 0$) or a parabolic advection-diffusion equation ($\gamma = 1$) when ε tends to 0.

In this paper, we will analyse the properties of the projective integration method in both the parabolic and the hyperbolic scaling. The numerical experiments in the present paper focus on the parabolic scaling, in which equation (3) becomes

$$\partial_t f^\varepsilon + \frac{v}{\varepsilon} \partial_x f^\varepsilon = \frac{\mathcal{M}_v(u^\varepsilon) - f^\varepsilon}{\varepsilon^2},$$

with macroscopic limit

$$\partial_t u^\varepsilon + \partial_x (A(u^\varepsilon)) = d \, \partial_{xx} u^\varepsilon. \tag{11}$$

Besides linear advection, we will also consider the viscous Burgers' equation, which is obtained when choosing $A(u) = u^2$. Numerical examples in the hyperbolic scaling are given in [43], which also discusses the generalisation to multiple space dimensions.

2.2   A kinetic semiconductor equation

While the numerical analysis of the presented algorithms is restricted to the above kinetic equation with $A(u)$ linear, we will also provide numerical results for a second model problem, in which macroscopic advection

4

does not originate from the Maxwellian in the collision operator, but from an external force field. To this end, we consider a kinetic equation that is inspired by the semiconductor equation [19],

$$\partial_t f^\varepsilon + \frac{1}{\varepsilon}\left(v\partial_x f^\varepsilon + F\partial_v f^\varepsilon\right) = \frac{u^\varepsilon - f^\varepsilon}{\varepsilon^2},$$
$$F = -\nabla \cdot \Phi, \qquad \Delta\Phi = u^\varepsilon. \tag{12}$$

This equation describes the evolution of the distribution function $f^\varepsilon(x, v, t)$, in which now an acceleration also appears due to an electric force F resulting from a coupled Poisson equation for the electric potential Φ. The velocity space is given by $V = \mathbb{R}$ endowed with the Gaussian measure $d\mu(v) = (2\pi)^{-1/2}\exp(-v^2/2)dv$.

3 High-order projective integration

The algorithm we propose in this paper is a high-order Runge–Kutta extension of the projective integration method [18, 37], which will turn out to be a fully explicit, arbitrary order, asymptotic-preserving time integration method for the kinetic equation (3). The asymptotic-preserving property [31] implies that, in the limit when ϵ tends to zero, an ε-independent time step constraint can be used, similar to the hyperbolic CFL-constraint for the limiting equation (10), depending on the scaling of (3). To achieve this, the projective integration method combines a few small time steps with a naive (*inner*) time-stepping method with a much larger (*projective, outer*) time step. The asymptotic-preserving property will then follow from the observation that both the size of the outer time step and the *number* of inner steps are independent of ε, resulting in a total computational cost that is independent of ε.

In sections 3.1 and 3.2, we discuss the inner and outer integrators, respectively. We then discuss the stability regions of the projective integration method in section 3.3.

3.1 Inner integrator

We intend to integrate (3) on a uniform, constant in time, periodic spatial mesh with spacing Δx, consisting of I mesh points $x_i = i\Delta x$, $0 \leq i \leq I$, with $I\Delta x = 1$, and a uniform time mesh with time step δt, i.e., $t^k = k\delta t$ and $k \geq 0$. The numerical solution on this mesh is denoted as $f_{i,j}^k$, where we have dropped the dependence on ε in the numerical solution for conciseness. After discretising in space, we obtain a semi-discrete system of ordinary differential equations

$$\dot{f} = \mathcal{D}_t(f), \qquad \mathcal{D}_t(f) := -\frac{1}{\varepsilon^\gamma}\mathcal{D}_{x,v}(f) + \frac{1}{\varepsilon^{\gamma+1}}\left(\mathcal{M}_v(u) - f\right), \tag{13}$$

where $\mathcal{D}_{x,v}(\cdot)$ represents a suitable discretisation of the first spatial derivative and $u = \langle f \rangle$. In the parabolic case, central differences are necessary (see [37] and the next sections) and in the related numerical experiments, we will use a fourth order discretisation,

$$(\mathcal{D}_{x,v_j}f)_{i,j} = \frac{-f_{i+2,j} + 8f_{i+1,j} - 8f_{i-1,j} + f_{i-2,j}}{12\Delta x}. \tag{14}$$

In the hyperbolic case, some type of upwinding needs to be performed and we will use, in the numerical experiments, a third order upwind biased scheme,

$$\begin{cases} (\mathcal{D}_{x,v_j}f)_{i,j} = v_j\dfrac{2f_{i+1,j} + 3f_{i,j} - 6f_{i-1,j} + f_{i-2,j}}{6\Delta x} & \text{if } v_j > 0, \\[3mm] (\mathcal{D}_{x,v_j}f)_{i,j} = v_j\dfrac{-f_{i+2,j} + 6f_{i+1,j} - 3f_{i,j} - f_{i-1,j}}{6\Delta x} & \text{if } v_j < 0. \end{cases} \tag{15}$$

Combined with a forward Euler time discretisation, we obtain

$$f_{i,j}^{k+1} = f_{i,j}^k - \frac{\delta t}{\varepsilon^\gamma}\mathcal{D}_{x,v_j}(f^k)_{i,j} + \frac{\delta t}{\varepsilon^{\gamma+1}}(\mathcal{M}_{v_j}(u_i^k) - f_{i,j}^k) \qquad \forall 1 \leq i \leq I, \quad 1 \leq j \leq J, \tag{16}$$

which we also denote using the shorthand notation

$$f^{k+1} = \mathcal{S}_{\delta t}(f^k), \qquad k = 0, 1, \ldots \tag{17}$$

In the context of projective integration, it does not make sense to investigate higher order methods for the inner integration. Some remarks on this fact are made in [43].

3.2 Outer integrator

The model problem we are dealing with is clearly stiff because of the presence of the small Knudsen parameter ε, leading to a time step restriction for the naive scheme (16) of $O(\varepsilon^{\gamma+1})$ due to the relaxation term. However, as ε goes to 0, we are able to obtain a limiting equation for which a standard finite volume/forward Euler method only needs to satisfy a stability restriction of the form $\Delta t \leq C\Delta x^{\gamma+1}$, with C a constant that depends on the specific choice of the scheme and the parameters of the equation.

In [37], the projective integration technique was proposed to accelerate brute force integration; the idea, originating from [18], is the following. Starting from a numerical solution f^N at time $t^N = N\Delta t$, one first takes $K+1$ *inner* steps of size δt, $f^{N,k+1} = S_{\delta t}(f^{N,k})$, $k = 0, \ldots, K$, in which the superscript pair (N, k) represents a numerical solution by means of the inner scheme at time $t^{N,k} = N\Delta t + k\delta t$. The aim is to obtain a discrete derivative to be used in the *outer* step to compute $f^{N+1} = f^{N+1,0}$ via extrapolation in time, e.g.,

$$f^{N+1} = f^{N,K+1} + (\Delta t - (K+1)\delta t) \frac{f^{N,K+1} - f^{N,K}}{\delta t}.$$

This method is called projective forward Euler (PFE), and it is the simplest instantiation of this class of integration methods [18, 37].

In this paper, we present a particular higher order extension of this idea, based on Runge–Kutta methods. Let us denote a general explicit S-stage Runge–Kutta method for equation (3) with time step Δt as

$$\begin{cases} f_s^{N+c_s} &= f^N + \Delta t \sum_{l=1}^{s-1} a_{s,l} k_l \\ k_s &= \mathcal{D}_t\left(f_s^{N+c_s}\right) \end{cases}, \qquad 1 \leq s \leq S,$$

$$f^{N+1} = f^N + \Delta t \sum_{s=1}^{S} b_s k_s,$$

with $\mathcal{D}_t$ defined in (13). As in [26], we call the matrix $A = (a_{s,l})_{s,l=2}^{S}$ the RK matrix, $b = (b_s)_{s=1}^{S}$ the RK weights and $c = (c_s)_{s=1}^{S}$ the RK nodes. The values k_s are called the RK stages, and represent an approximation of the time derivative at time $t = t^N + c_s\Delta t$. The weights b_s and c_s are chosen simultaneously, and correspond to a Gauss quadrature approximation of the integration from $t = t^N$ to t^{N+1}. To ensure consistency, these coefficients satisfy the following assumptions (see, e.g., [26]):

Assumption 3.1 (Runge–Kutta coefficients)**.** *The Runge–Kutta coefficients satisfy* $0 \leq c_s \leq 1$, *resp.* $0 \leq b_s \leq 1$, *and*

$$\sum_{s=1}^{S} b_s = 1, \qquad \sum_{l=1}^{S-1} a_{s,l} = c_s, \quad 1 \leq s \leq S.$$

(Note that these assumptions imply that $c_1 = 0$ by the convention that $\sum_1^0 \cdot = 0$).

In the higher order projective integration method, we proceed, by analogy with the projective forward Euler method, by replacing each time derivative evaluation k_s by $K+1$ steps of an inner integrator and a

$$
\begin{array}{c|c}
c & A \\
\hline
 & b^T
\end{array}
\qquad\qquad
\begin{array}{c|cc}
0 & & \\
1/2 & 1/2 & \\
\hline
 & 0 & 1
\end{array}
\qquad\qquad
\begin{array}{c|cccc}
0 & & & & \\
1/2 & 1/2 & & & \\
1/2 & 0 & 1/2 & & \\
1 & 0 & 0 & 1 & \\
\hline
 & 1/6 & 1/3 & 1/3 & 1/6
\end{array}
$$

Figure 2: Butcher tableaux for Runge-Kutta methods. Left: general notation; middle: RK2 method (second order); right: RK4 method (fourth order).

time derivative estimate as follows (with $f^{N,0} = f^N$ for consistency):

$$
s = 1 : \begin{cases}
f^{N,k} & = f^{N,k-1} + \delta t \mathcal{D}_t(f^{N,k-1}), \qquad 1 \leq k \leq K + 1 \\
k_1 & = \dfrac{f^{N,K+1} - f^{N,K}}{\delta t},
\end{cases}
\tag{18}
$$

$$
2 \leq s \leq S : \begin{cases}
f_s^{N+c_s} & = f^{N,K+1} + (c_s \Delta t - (K+1)\delta t) \sum_{l=1}^{s-1} \dfrac{a_{s,l}}{c_s} k_l, \\
f^{N+c_s,k} & = f^{N+c_s,k-1} + \delta t \mathcal{D}_t(f^{N+c_s,k-1}), \qquad 1 \leq k \leq K + 1 \\
k_s & = \dfrac{f^{N+c_s,K+1} - f^{N+c_s,K}}{\delta t},
\end{cases}
\tag{19}
$$

$$
f^{N+1} = f^{N,K+1} + (\Delta t - (K+1)\delta t) \sum_{s=1}^{S} b_s k_s.
\tag{20}
$$

In the following sections, it will be shown that the small time step should be taken as

$$
\delta t = \varepsilon^{\gamma+1}.
$$

Note that the stages k_s now record a finite difference approximation of the time derivative at time $t = t^N + c_s \Delta t + (K+1)\delta t$, not at time $t = t^N + c_s \Delta t$. Hence, one should, in principle, adjust the weight b_s to keep the Gaussian quadrature interpretation of the Runge–Kutta method, see, e.g., [39] for work in this direction. However, as will be shown in section 5, this additional consistency error will be negligible in the limit when ε tends to 0, which is the relevant limit in this paper.

In the numerical experiments, we will specifically use the projective Runge–Kutta methods of orders 2 and 4 represented by the Butcher tableaux in Figure 2.

3.3 Stability of higher order projective integration

Let us now study the linear stability regions of the higher order Runge–Kutta projective integration methods that were devised above. As is traditional, we introduce to this end the Dahlquist test equation and its corresponding inner integrator,

$$
\dot{y} = \lambda y, \qquad y^{k+1} = \tau(\lambda \delta t) y^k, \qquad \lambda < 0.
\tag{21}
$$

As in [18], we call $\tau(\lambda \delta t)$ the *amplification factor* of the inner integrator. (For instance, if the inner integrator is forward Euler, we have $\tau(\lambda \delta t) = 1 + \lambda \delta t$.) The inner integrator is linearly stable if $|\tau| \leq 1$. The analysis below will reveal for which values of τ the projective integration method is also stable. In section 4.3, this analysis will be combined with an analysis of the spectrum of the kinetic equation (3) to determine the method parameters δt, Δt and K of the projective integration method.

A projective Runge–Kutta method applied to (21) can be written as

$$
y^{N+1} = \sigma(\tau; \Delta t, \delta t, K) y^N,
\tag{22}
$$

7

which is stable when $|\sigma(\tau; \Delta t, \delta t, K)| \leq 1$. For projective forward Euler, we have

$$\sigma^{PFE}(\tau; \Delta t, \delta t, K) = \left[\left(\frac{\Delta t - (K+1)\delta t}{\delta t} + 1 \right) \tau - \frac{\Delta t - (K+1)\delta t}{\delta t} \right] \tau^K. \tag{23}$$

Given the kinetic equation (3), the goal in this paper is to take a projective time step $\Delta t = O(\Delta x^{\gamma+1})$, whereas $\delta t = O(\varepsilon^{\gamma+1})$ necessarily to ensure stability of the inner brute-force forward Euler integration. Since we are interested in the limit $\varepsilon \to 0$ for fixed Δx, we therefore look at the limiting stability regions as $\Delta t / \delta t \to \infty$. In this regime, it is shown in [18] that the values τ for which the condition (23) is satisfied lie in the union of two separated disks $\mathcal{D}_1^{PFE} \cup \mathcal{D}_2^{PFE}$ where

$$\mathcal{D}_1^{PFE} = \mathcal{D}\left(1 - \frac{\delta t}{\Delta t}, \frac{\delta t}{\Delta t} \right) \text{ and } \mathcal{D}_2^{PFE} = \mathcal{D}\left(0, \left(\frac{\delta t}{\Delta t} \right)^{1/K} \right). \tag{24}$$

The eigenvalues in $\mathcal{D}_2^{PFE}$ correspond to modes that are quickly damped by the time-stepper, whereas the eigenvalues in $\mathcal{D}_1^{PFE}$ correspond to slowly decaying modes. When the method parameters δt, Δt and K are suitably chosen, the projective integration method then allows for accurate integration of the modes in $\mathcal{D}_1^{PFE}$ while maintaining stability for the modes in $\mathcal{D}_2^{PFE}$.

We now show how the stability regions of higher order projective Runge–Kutta schemes relate to those of projective forward Euler when $\delta t / \Delta t$ tends to 0.

Theorem 3.2 (Stability of higher order projective Runge–Kutta methods). *Assume the inner integrator is stable, i.e., $|\tau| \leq 1$, and δt, K and Δt are chosen such that the projective forward Euler method is stable. Then, a projective Runge–Kutta method is also stable if it satisfies Assumptions 3.1 and the convexity condition*

$$0 \leq a_{s,l} \leq c_s, \qquad \forall 1 \leq l \leq s, \ 1 \leq s \leq S. \tag{25}$$

Such a result is classical for regular Runge–Kutta methods [26]. Here, however, we also provide the proof in the *projective* Runge–Kutta case, to show that the above property holds both for the stability domain corresponding to slow eigenvalues and for the stability domain corresponding to quickly damped eigenvalues.

Proof. Let us first introduce, as in [18], $M = (\Delta t - (K+1)\delta t)/\delta t$ and, similarly, $M_s = (c_s\Delta t - (K+1)\delta t)/\delta t$, and remark that $M_s \leq c_s M \leq M$ is satisfied for all $s \in [1, S]$. We can then rewrite the Runge–Kutta scheme (18)-(19)-(20) for the test equation (21):

$$\begin{cases} k_1 & =: \kappa_1(\tau)y^N = \dfrac{\tau^{K+1} - \tau^K}{\delta t} y^N \\ k_s & =: \kappa_s(\tau)y^N = \dfrac{\tau^{K+1} - \tau^K}{\delta t} \left(\tau^{K+1} + (M_s\delta t) \sum_{l=1}^{s-1} \dfrac{a_{s,l}}{c_s} \kappa_l \right) y^N, \qquad 2 \leq s \leq S, \\ y^{N+1} & =: \sigma(\tau)y^N = \left(\tau^{K+1} + (M\delta t) \sum_{s=1}^{S} b_s \kappa_s \right) y^N, \end{cases}$$

where we have suppressed the dependence of κ and σ on K, δt and Δt but emphasized the dependence on τ.

The proof then amounts to showing that the condition $|\sigma| \leq 1$ is satisfied as soon as the stability condition for the projective forward Euler scheme, i.e.,

$$\left| \left((M+1)\tau - M \right) \tau^K \right| \leq 1, \tag{26}$$

is satisfied. The proof is split up in three steps:

- We first remark that if condition (26) is satisfied, this implies that

$$\left| \left((\alpha M + 1)\tau - \alpha M \right) \tau^K \right| \leq 1, \tag{27}$$

for all $\alpha \in [0, 1]$, since $\alpha M/(\alpha M + 1) \leq M/(M + 1)$, so that $\mathcal{D}(M/(M + 1)) \subset \mathcal{D}(\alpha M/(\alpha M + 1))$.

8

- Next, we prove by induction that

$$\kappa_s \leq \frac{\left| \tau^{K+1} - \tau^K \right|}{\delta t}, \qquad 1 \leq s \leq S. \tag{28}$$

Clearly, this statement is true for $s = 1$. For $s > 1$, we have

$$\kappa_s = \frac{\tau^{K+1} - \tau^K}{\delta t} \left(\tau^{K+1} + (M_s \delta t) \sum_{l=1}^{s-1} \frac{a_{s,l}}{c_s} \kappa_l \right).$$

Assume that for $n \in \{1, \dots, s-1\}$, $s \geq 2$:

$$|\kappa_n| \leq \frac{\tau^{K+1} - \tau^K}{\delta t}. \tag{29}$$

We thus need to show that

$$\left| \tau^{K+1} + (M_s \delta t) \sum_{l=1}^{s-1} \frac{a_{s,l}}{c_s} \kappa_l \right| \leq 1.$$

To this end, we write

$$\tau^{K+1} + (M_s \delta t) \sum_{l=1}^{s-1} \frac{a_{s,l}}{c_s} \kappa_l = \tau^{K+1} + M_s \left(\tau^{K+1} - \tau^K \right) \sum_{l=1}^{s-1} \frac{a_{s,l}}{c_s} \frac{\kappa_l \delta t}{\tau^{K+1} - \tau^K}$$

$$= ((\alpha M_s + 1)\tau - \alpha M_s) \tau^K,$$

with

$$\alpha = \frac{M_s}{M} \sum_{l=1}^{s-1} \frac{a_{s,l}}{c_s} \frac{\kappa_l \delta t}{\tau^{K+1} - \tau^K}.$$

Using (25), the induction hypothesis (29) and the fact that $M_s \leq M$, we deduce that $0 \leq \alpha \leq 1$, from which, using (27), we conclude (28).

- Now we are ready to show that (3.3) holds, since the latest result is valid for $s = S$. Using the same reasoning, we can rewrite σ as:

$$\sigma = ((\beta M + 1)\tau - \beta M) \tau^k, \qquad \beta = \sum_{s=1}^{S} b_s \kappa_s \frac{\delta t}{\tau^{K+1} - \tau^K},$$

from which, using (28), $0 \leq \beta \leq 1$ and assumptions 3.1, we deduce (3.3).

$\qquad \square$

As for projective forward Euler, the stability region breaks up into two parts when $\delta t / \Delta t$ tends to 0. By performing an asymptotic expansion of σ (see (22)) in terms of $\delta t / \Delta t$, we can obtain a parameterisation of the boundary of both regions, defined by the set of values τ for which $|\sigma(\tau; \Delta t, K, \delta t)| = 1$. We have the following result:

Proposition 3.3. *In the limit when $\delta t / \Delta t$ tends to 0, the stability region of a projective Runge–Kutta method consists of two regions $\mathcal{R}_1^{\mathrm{PRK}} \cup \mathcal{R}_2^{\mathrm{PRK}}$. The boundary of $\mathcal{R}_1^{\mathrm{PRK}}$ is given by an asymptotic expansion of the form*

$$\tau(\theta) = 1 + C_1(\theta) \left(\frac{\delta t}{\Delta t} \right) + C_2(\theta) \left(\frac{\delta t}{\Delta t} \right)^2 + \ h.o.t., \qquad 0 \leq \theta \leq 2\pi,$$

whereras the boundary of $\mathcal{R}_2^{\mathrm{PRK}}$ can be expanded as

$$\tau(\theta) = C_1'(\theta) \left(\frac{\delta t}{\Delta t} \right)^{1/K} + C_2'(\theta) \left(\frac{\delta t}{\Delta t} \right)^{2/K} + \ h.o.t., \qquad 0 \leq \theta \leq 2\pi.$$

The proof, containing also the expressions for $C_{1,2}(\theta)$ and $C'_{1,2}(\theta)$, is given in the Appendix, which also contains the expressions of the projective Runge–Kutta methods with Butcher tableaux in Figure 2. An additional observation, which we will state here without proof, is that in the limit when $\delta t/\Delta t$ tends to 0, the stability regions of lower order methods are contained within those of higher-order methods, i.e., the stability regions satisfy

$$\mathcal{R}_1^{PRK,p+1} \supseteq \mathcal{R}_1^{PRKp} \supseteq \mathcal{D}_1^{PFE} \quad \text{and} \quad \mathcal{R}_2^{PRKp+1} \supseteq \mathcal{R}_2^{PRKp} \supseteq \mathcal{D}_2^{PFE}, \qquad \forall p \geq 1,$$

in which the integer p indicates the order of the method.

We illustrate the shape of these stability domains for the classical second-order and fourth order Runge–Kutta method whose tableaux are given in figure 2. The stability regions are shown in figure 3. The figure illustrates theorem 3.2, and additionally shows that the stability regions scale with $\Delta t/\delta t$ in the same way as for the projective forward Euler method. The shape of the stability regions, however, depends on the method used. It can be checked that the region $\mathcal{R}_1^{PRK}$ converges to the stability domain of the corresponding classical Runge–Kutta method when $\delta t/\Delta t$ tends to 0.
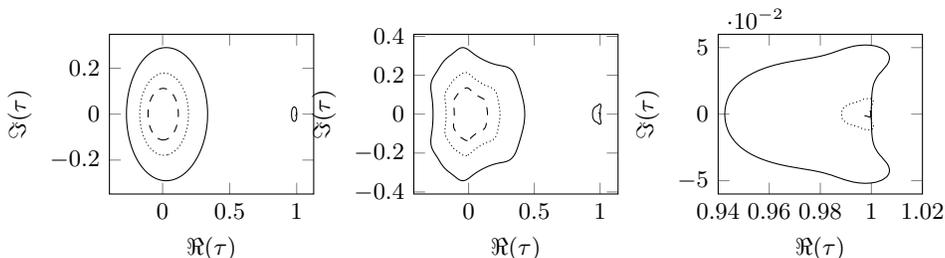

Figure 3: The two leftmost pictures show (respectively) the stability regions for the PRK2 method and the PRK4 method, while the picture on the right shows a zoom on the region of the PRK4 method near 1. Parameters: $\Delta t = 1 \times 10^{-3}, K = 3$ and $\delta t = 1 \times 10^{-6}$ (dashed), $\delta t = 4 \times 10^{-6}$ (dotted) and $\delta t = 1.6 \times 10^{-5}$.

The main conclusion of the above analysis is that, whereas the stability regions of higher order projective Runge–Kutta methods differ from those of projective forward Euler in their precise shape, their qualitative dependence on the parameters of projective integration (δt, K and Δt) is the same, and method parameters that are suitable for projective forward Euler, will also be suitable for the higher order projective Runge–Kutta method.

4   Stability analysis

We are now ready to study the stability of the projective integration schemes for the kinetic equation (3). After introducing some notation in section 4.1, we compute bounds on the spectrum of the inner integrator (16) with a linear Maxwellian (7) with $A(u) = u$ in section 4.2. Subsequently, we look into suitable parameter choices for the projective integration schemes (section 4.3).

4.1   Notation and assumptions

We first rewrite the semi-discretized kinetic equation (13) in the (spatial) Fourier domain,

$$\partial_t \hat{F}(\zeta) = B\,\hat{F}(\zeta), \qquad \text{with} \qquad B = \frac{1}{\varepsilon^{\gamma+1}}\left(\varepsilon D + MP - \mathcal{I}\right), \tag{30}$$

with $\hat{F} \in \mathbb{R}^J$, the matrices B, M, $P \in \mathbb{R}^{J \times J}$, and $\mathcal{I}$ the identity matrix of dimension J. In (30), the matrix D represents minus the (diagonal) Fourier matrix of the spatial discretisation chosen for the convection part,

P is the rank 1 Fourier matrix of the averaging of f over all velocities,

$$P := ee^T, \qquad e = \frac{1}{\sqrt{J}}(1, \ldots, 1)^T \in \mathbb{R}^J,$$

and the invertible matrix M represents the Fourier transform of the Maxwellian, $M = \mathcal{I} + \varepsilon^\gamma V^{-1}$, with V the diagonal matrix with elements v_j given in (5). For the spatial discretisations in equations (14) and (15), the matrix D is

$$\begin{cases} D = -\imath \dfrac{8\sin(\zeta) - \sin(2\zeta)}{6\Delta x} V & \text{(parabolic)}, \\[4mm] D = -\left(\dfrac{3 - 4\cos(\zeta) + \cos(2\zeta)}{6\Delta x}\right) V - \imath \left(\dfrac{8\sin(\zeta) - \sin(2\zeta)}{6\Delta x}\right) V & \text{(hyperbolic)}. \end{cases}$$

From now on, we write $D_j = \alpha_j + \imath\beta_j$ for $j \in \{1, \ldots, J\}$. Thus, we have

$$\alpha_j = -\frac{|v_j|}{6\Delta x}(3 - 4\cos(\zeta) + \cos(2\zeta)) \qquad \beta_j = -\frac{v_j}{6\Delta x}(8\sin(\zeta) - \sin(2\zeta)), \tag{31}$$

for the third order upwind scheme, whereas

$$\alpha_j = 0 \qquad \beta_j = -v_j \frac{8\sin(\zeta) - \sin(2\zeta)}{6\Delta x}. \tag{32}$$

for the fourth order central scheme. We also define

$$\tilde{e} = Me = (\mathcal{I} + \varepsilon^\gamma V^{-1})e, \tag{33}$$

from which we obtain $MP = \tilde{e}e^T$. We write the Fourier transform of (16) as

$$\hat{F}^{k+1} = S_{\delta t}\, \hat{F}^k = (\mathcal{I} + \delta t B)\, \hat{F}^k = \left(1 - \frac{\delta t}{\varepsilon^{\gamma+1}}\right)\mathcal{I} + \frac{\delta t}{\varepsilon^{\gamma+1}} A, \tag{34}$$

where A is defined as $MP + \varepsilon D$. It is clear that the amplification factors $\tau = (\tau_1, \ldots, \tau_J)$ of the forward Euler scheme (which are the eigenvalues of $S_{\delta t}$) and the eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_J)$ of the matrix A are related via

$$\tau_j = \left(1 - \frac{\delta t}{\varepsilon^{\gamma+1}}\right) + \frac{\delta t}{\varepsilon^{\gamma+1}}\lambda_j, \qquad 1 \le j \le J.$$

To locate the spectrum, we assume the velocity space is symmetric (see (5)),

$$v_{J-j} = -v_j \qquad 1 \le j \le J/2, \tag{35}$$

so that

$$D_{J-j} = \overline{D}_j \qquad 1 \le j \le J/2.$$

4.2 Spectrum of the inner integrators

We have the following result for the spectrum of $A = MP + \varepsilon D$.

Theorem 4.1. *Under the assumptions in section 4.1, the spectrum of the matrix $A = MP + \varepsilon D$ satisfies*

$$\mathrm{Sp}(A) \subset \left(\mathcal{D}\left(0, \varepsilon C \max_{j \in \mathcal{J}^+}(|\alpha_j| + |\beta_j|)\right)\right) \cup \{\lambda(\varepsilon)\}$$

where the constant C depends on the parameters $(\alpha_j)_{j=1}^J$ and $(\beta_j)_{j=1}^J$ of the spatial discretisation scheme and the chosen velocities $(v_j)_{j=1}^J$. The dominant eigenvalue $\lambda(\varepsilon)$ is simple and can be expanded as

$$
\mathrm{Re}(\lambda(\varepsilon)) = 1 + \varepsilon\langle\alpha\rangle + \varepsilon^2\left(\langle((\langle\alpha\rangle - \alpha)^2\rangle - \langle\beta^2\rangle + \delta_\gamma\left\langle\frac{\beta}{v}\right\rangle^2\right) + o(\varepsilon^2)
$$

$$
\mathrm{Im}(\lambda(\varepsilon)) = \varepsilon\delta_\gamma\left\langle\frac{\beta}{v}\right\rangle + \varepsilon^2\left(\delta_\gamma\left\langle\left(\left\langle\frac{\beta}{v}\right\rangle - \beta\right)(\langle\alpha\rangle - \alpha)\right\rangle + \delta_{\gamma-1}\left\langle\frac{\beta}{v}\right\rangle\right) + o(\varepsilon^2),
$$

where we used δ_γ in the sense of the classical Kronecker delta symbol, where $\delta_\gamma = 1$ if $\gamma = 0$ and zero otherwise.

The proof of theorem 4.1 has the same structure as the proof in [37]. However, due to the presence of the Maxwellian $\mathcal{M}_v$, each of the intermediate steps becomes more involved. We split up these steps in several lemmas.

Lemma 4.2. *The rank-one matrix MP is a projection matrix.*

Proof. We need to show that $(MP)^2 = MP$. Using the definitions introduced above, we get

$$
\begin{aligned}
(MP)^2 &= (\mathcal{I} + \varepsilon^\gamma V^{-1})(ee^T)^2 + \varepsilon^\gamma(\mathcal{I} + \varepsilon^\gamma V^{-1})ee^T V^{-1}ee^T \\
&= (\mathcal{I} + \varepsilon^\gamma V^{-1})ee^T = MP,
\end{aligned}
$$

where, in the last line, we have used (i) the fact that $e^T V^{-1} = 0$ to eliminate the second term (since the velocity space is assumed to be odd), and (ii) $e^T e = 1$, from which we obtain $(ee^T)^2 = ee^T$. $\square$

The following corollary is an immediate consequence.

Corollary 4.3. *The matrix MP has one eigenvalue $\lambda_1 = 1$ and all other eigenvalues vanish, i. e. $\lambda_j = 0$, $2 \leq j \leq J$.*

Lemma 4.4. *Consider the matrix $A = MP + \tilde{D}$, and assume*

$$
\tilde{D} = \mathrm{diag}(\tilde{D}_1, \ldots, \tilde{D}_J), \quad \text{where } \tilde{D}_j = \tilde{D}_{j'} \text{ implies } j = j'. \tag{38}
$$

Then, the eigenspaces of A are of dimension 1 and no $\tilde{D}_j, 1 \leq j \leq J$ is an eigenvalue of A.

Proof. Let (λ, W) be an eigenvalue and an associated eigenvector of A. This implies

$$
\left(MP + \tilde{D}\right)W = \lambda W \langle W\rangle\tilde{e} + \tilde{D}W = \lambda W, \quad \text{with } \langle W\rangle = e^T W. \tag{39}
$$

Assume now $\langle W\rangle = 0$, from which we infer that $\tilde{D}W = \lambda W$. Since $\langle W\rangle = 0$ with $W \neq 0$, there exists at least two indices j_1 and j_2 such that $W_{j_1}, Wj_2 \neq 0$. However, this implies that $\lambda = D_{j_1} = D_{j_2}$, which violates assumption (38).

So necessarily $\langle W\rangle \neq 0$. Then (39) implies $W = \left(\lambda\mathcal{I} - \tilde{D}\right)^{-1}\langle W\rangle\tilde{e}$ that is, all the eigenspaces are of dimension 1 and no D_j can be an eigenvalue of A. $\square$

Let us, from now on, choose $\tilde{D} = \varepsilon D$, and investigate the matrix $A = MP + \varepsilon D$.

Lemma 4.5. *Introducing $Q(\lambda) := \prod_{j=1}^J(\varepsilon D_j - \lambda)$, the characteristic polynomial $\chi_A(\lambda)$ of $A = MP + \varepsilon D$ can be written as*

$$
\chi_A(\lambda) = Q(\lambda)\left(1 - \frac{1}{J}\sum_{j=1}^J\frac{1 + \varepsilon^\gamma/v_j}{\lambda - \varepsilon D_j}\right) \quad .
$$

Proof. We start by writing

$$
\chi_A(\lambda) =
\begin{vmatrix}
\tilde{e}_1/\sqrt{J} + \varepsilon D_1 - \lambda & \tilde{e}_1/\sqrt{J} & \cdots & \tilde{e}_1\sqrt{J} \\
\tilde{e}_2/\sqrt{J} & \tilde{e}_2/\sqrt{J} + \varepsilon D_2 - \lambda & \cdots & \tilde{e}_2/\sqrt{J} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{e}_J/\sqrt{J} & \tilde{e}_J/\sqrt{J} & \cdots & \tilde{e}_J/\sqrt{J} + \varepsilon D_J - \lambda
\end{vmatrix},
$$

with $\tilde{e}_j = (1/\sqrt{J})\,(1 + \varepsilon^\gamma/v_j)$. This is the determinant of an arrow matrix

$$
a =
\begin{pmatrix}
d_1 & r_2 & & \cdots & r_J \\
c_2 & d_2 & 0 & & \\
c_3 & 0 & d_3 & 0 & \\
\vdots & & & \ddots & \\
c_J & & & & d_J
\end{pmatrix},
$$

the determinant of which is

$$
\det(a) = \prod_{j=1}^{J} d_j - \sum_{j=2}^{J} \left(c_j r_j \prod_{j'=2, j' \neq j}^{J} d_{j'} \right).
$$

So, after identifying $d_1 = \tilde{e}_1/\sqrt{J} + \varepsilon D_1 - \lambda$ and $d_j = \varepsilon D_j - \lambda$, $c_j = \tilde{e}_j/\sqrt{J}$ and $r_j = -(\varepsilon D_1 - \lambda)$, for $2 \leq j \leq J$ some elementary manipulations yield $\chi_A(\lambda)$

$$
\chi_A(\lambda) = Q(\lambda) \left(1 - \frac{1}{J} \sum_{j=1}^{J} \frac{1 + \varepsilon^\gamma/v_j}{\lambda - \varepsilon D_j} \right),
$$

where we made use of (33). This concludes the proof. $\qquad\square$

To prove theorem 4.1, we will consider the characteristic polynomial $\chi_A(\lambda)$ to be a perturbation of the characteristic polynomial that was studied in [37]. We recall the following theorem from [37] in the notation of the present paper.

Theorem 4.6 (Proposition 4.1 in [37])**.** *Consider the matrix $A_0 = P + \varepsilon D$, assuming (38) and (35). Then, the corresponding characteristic polynomial is*

$$
\chi_{A_0}(\lambda) = Q(\lambda) \left(1 - \frac{1}{J} \sum_{j=1}^{J} \frac{1}{\lambda - \varepsilon D_j} \right)
$$

and its eigenvalues satisfy

$$
\mathrm{Sp}(A) \subset \left(\mathcal{D}\left(0, \frac{\varepsilon}{J} \max_{j \in \mathcal{J}^+} (|\alpha_j| + |\beta_j|) \right) \right) \cup \{\lambda(\varepsilon)\}
$$

where the real eigenvalue $\lambda(\varepsilon)$ is simple and can be expanded as

$$
\lambda(\varepsilon) = 1 - \varepsilon \frac{\langle \alpha \rangle}{J} - \frac{\varepsilon^2}{J^2} \langle (\alpha - \langle \alpha \rangle)^2 + \beta^2 \rangle + o(\varepsilon^2).
$$

Since we know how to localize the roots of χ_{A_0}, we can use Rouché's theorem [51] to bound the eigenvalues of χ_A.

Proposition 4.7 (Rouché's theorem)**.** *If there exists a closed simple contour ζ in $\mathbb{C}$ encircling a compact $\mathcal{C}$, such that*

$$\forall \lambda \in \zeta, \chi_{A_0}(\lambda) \neq 0 \quad and \quad |\chi_A(\lambda) - \chi_{A_0}(\lambda)| < |\chi_{A_0}(\lambda)|, \tag{40}$$

then χ_A and χ_{A_0} have exactly the same number of roots in $\mathcal{C}$.

Everything is now in place to prove theorem 4.1.

Proof of theorem 4.1. The proof consists of two steps. First, we will construct, using Rouché's theorem, contours in which the eigenvalues of χ_A are known to be localized. In a second step, we will provide an asymptotic expansion for the dominant eigenvalue.

Step (i): Localization of eigenvalues   We start by writing

$$\chi_A(\lambda) = \chi_{A_0}(\lambda) + \frac{\varepsilon^\gamma}{J} \sum_{j=1}^{J} \frac{R_j(\lambda)}{v_j},$$

and aim at applying Rouché's theorem. We thus study the rational function

$$\mathcal{F} : \lambda \mapsto \frac{\chi_A(\lambda) - \chi_{A_0}(\lambda)}{\chi_{A_0}(\lambda)} = -\frac{\dfrac{\varepsilon^\gamma}{J} \sum_{j=1}^{J} \dfrac{1}{v_j} \dfrac{1}{\lambda - \varepsilon D_j}}{1 - \dfrac{1}{J} \sum_{j=1}^{J} \dfrac{1}{\lambda - \varepsilon D_j}}$$

and look for contours that contain the eigenvalues of χ_{A_0} and for which $|\mathcal{F}(\zeta)| < 1$.

- Let us first consider the dominant eigenvalue by enclosing the dominant eigenvalues of χ_{A_0} in a circle around $\lambda = 1$. To this end, we search a value of $r > 0$ such that (40) is satisfied on $\zeta = \{1 + \varepsilon^{\gamma+1} r e^{i\theta}, \theta \in [0, 2\pi)\}$. Performing a Taylor expansion of $1/(1 + \varepsilon^{\gamma+1} r e^{i\theta} - \varepsilon D_j)$ in terms of ε yields

$$
\begin{aligned}
\mathcal{F}(1 + \varepsilon^{\gamma+1} r e^{i\theta}) &= \frac{-\dfrac{\varepsilon^\gamma}{J} \sum_{j=1}^{J} \dfrac{1}{v_j} \left(\varepsilon D_j - \varepsilon^{\gamma+1} r e^{i\theta} \right) + O(\varepsilon^2)}{1 - \dfrac{1}{J} \sum_{j=1}^{J} 1 - \varepsilon^{\gamma+1} r e^{i\theta} + \varepsilon D_j + O(\varepsilon^2)} \\[2mm]
&= \frac{-\dfrac{\varepsilon^\gamma}{J} \sum_{j=1}^{J} \dfrac{D_j}{v_j} + O(\varepsilon)}{\varepsilon^\gamma r e^{i\theta} - \dfrac{1}{J} \sum_{j=1}^{J} D_j + O(\varepsilon)}
\end{aligned}
$$

where we have used the fact that $\frac{1}{J} \sum_j v_j = 0$. When choosing r such that $r > \frac{1}{J} \left| \sum_{j=1}^{J} D_j \right| + \frac{2}{J} \left| \sum_{j=1}^{J} \frac{D_j}{v_j} \right|$, we ensure that $|\mathcal{F}(\lambda)| < 1/2 + O(\varepsilon)$, from which one can conclude that χ_A and χ_{A_0} have the same number of zeroes, that is, 1, around $\lambda = 1$ in a neighbourhood of size $\varepsilon^{\gamma+1}$.

- Let us now consider the $J - 1$ remaining eigenvalues by considering the region around $\lambda = 0$. Again, we will make use of Rouché's theorem: let us find $r > 0$ such that (40) is satisfied on $\zeta = \{\varepsilon r e^{i\theta}, \theta \in [0, 2\pi)\}$. We thus study

$$\mathcal{F}(\varepsilon r e^{i\theta}) = -(\varepsilon^\gamma) \frac{\phi(r)}{\varepsilon - \psi(r)}$$

14

with

$$\psi : r \mapsto \frac{1}{J} \sum_{j=1}^{J} \frac{1}{re^{\imath\theta} - D_j}, \qquad \phi(r) : r \mapsto \frac{1}{J} \sum_{j=1}^{J} \frac{1}{v_j} \frac{1}{re^{\imath\theta} - D_j}.$$

Performing a Taylor expansion of $\phi(r)$ and $\psi(r)$ in $1/r$ yields:

$$\psi(r) = \frac{e^{-\imath\theta}}{r} + O\left(\frac{1}{r^2}\right), \quad \phi(r) = \frac{e^{-2\imath\theta}}{Jr^2} \sum_{j=1}^{J} \frac{D_j}{v_j} + O\left(\frac{1}{r^3}\right),$$

so

$$\mathcal{F}(\varepsilon re^{\imath\theta}) = -(\varepsilon^\gamma) \frac{e^{-\imath 2\theta}}{r^2} \frac{\frac{1}{J} \sum_{j=1}^{J} \frac{D_j}{v_j} + O\left(\frac{1}{r}\right)}{\varepsilon - \frac{e^{-\imath\theta}}{r} + O\left(\frac{1}{r^2}\right)} = (\varepsilon^\gamma) \frac{e^{-\imath\theta}}{r} \frac{1}{J} \sum_{j=1}^{J} \frac{D_j}{v_j} (1 + \varepsilon re^{\imath\theta} + O(\varepsilon^2)).$$

Choose $r > 2 \max\left(\frac{\varepsilon^\gamma}{J} \left|\sum_{j=1}^{J} \frac{D_j}{v_j}\right|, 1\right)$ to ensure that the main term in the Taylor expansion is in modulus less than $1/2$. Thus we can conclude that there are exactly $J - 1$ eigenvalues in a neighbourhood of $\lambda = 0$ of size ε.

Step (ii): Asymptotic expansion of dominant eigenvalue To obtain an asymptotic expansion of the dominant eigenvalue, we first define

$$S_j := \frac{1 + \varepsilon^\gamma/v_j}{\lambda - \varepsilon D_j} + \frac{1 - \varepsilon^\gamma/v_j}{\lambda - \varepsilon \bar{D}_j}, \quad \text{as well as} \quad \Sigma(\lambda(\varepsilon)) := \frac{1}{J} \sum_{j=1}^{J/2} S_j.$$

Now, given that $\lambda(\varepsilon)$ (close to 1) is a root of the characteristic polynomial $\chi_A(\lambda)$, we have $\Sigma(\lambda(\varepsilon)) = 1$. We therefore perform a Taylor expansion of $\Sigma(\lambda(\varepsilon))$ around $\varepsilon = 0$. We split λ in its real and imaginary part: $\lambda(\varepsilon) = x(\varepsilon) + \imath\, y(\varepsilon)$ and proceed by requiring (up to second order)

$$\Sigma(\lambda(\epsilon)) \equiv \Sigma(\varepsilon) = \Sigma(0) + \varepsilon \Sigma'(0) + \frac{\varepsilon^2}{2} \Sigma''(0) = 1.$$

Matching, for all powers of ε the real and imaginary parts of the left and right hand side, yields the conditions: $\text{Re}(\Sigma(0)) = 1, \text{Im}(\Sigma(0)) = 0$ and $\text{Re}(\Sigma^{(j)}(0)) = \text{Im}(\Sigma^{(j)}(0)) = 0 \quad \forall j \geq 1$. From these conditions, asymptotic expansions of $x(\varepsilon)$ and $y(\varepsilon)$ in terms of ε can be derived. Let us write

$$x(\varepsilon) = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + O(\varepsilon^3), \qquad y(\varepsilon) = y_0 + \varepsilon y_1 + \varepsilon^2 y_2 + O(\varepsilon^3)$$

Then, we get, for the zeroth order term,

$$\frac{2}{J} \sum_{j=1}^{J/2} \frac{x_0}{x_0^2 + y_0^2} = 1 \qquad -\frac{2}{J} \sum_{j=1}^{J/2} \frac{y_0 \left(1 + \varepsilon^\gamma/v_j\right)}{x_0^2 + y_0^2} = 0,$$

which implies that $x_0 = 1$ and $y_0 = 0$. (This is consistent with the derivation based on Rouché's theorem above.)

Next, we determine the terms of order ε,

$$\text{Re}(\Sigma)(0) = -\frac{2}{J} \sum_{j=1}^{J/2} x_1 - \alpha_j \qquad \text{Im}(\Sigma)(0) = \frac{2}{J} \sum_{j=1}^{J/2} \left(-y_1 + \delta_\gamma \frac{\beta_j}{v_j}\right),$$

15

from which we conclude that $x_1 = \langle \alpha \rangle$ and $y_1 = \delta_\gamma \langle \beta/v \rangle$. Finally, for the second order terms, we find

$$
\begin{cases}
\displaystyle\sum_{j=1}^{J/2} -2x_2 + 4(\langle\alpha\rangle - \alpha_j)^2 - 4(y_1^2 + \beta_j^2) + 8\delta_\gamma y_1 \frac{\beta_j}{v_j} = 0 \\
\displaystyle\sum_{j=1}^{J/2} -2y_2 + 4\delta_\gamma \left(\delta_\gamma\left\langle\frac{\beta}{v}\right\rangle - \beta_j\right)(\langle\alpha\rangle - \alpha_j) + \delta_{\gamma-1}\frac{\beta_j}{v_j} = 0,
\end{cases}
$$

from which we conclude that

$$
\begin{aligned}
x_2 &= 2\left\langle (\langle\alpha\rangle - \alpha)^2 - \beta^2 + \delta_\gamma\left\langle\frac{\beta}{v}\right\rangle^2 \right\rangle \\
y_2 &= 2\delta_\gamma\left\langle \left(\left\langle\frac{\beta}{v}\right\rangle - \beta\right)(\langle\alpha\rangle - \alpha) \right\rangle + 2\delta_{\gamma-1}\left\langle\frac{\beta}{v}\right\rangle
\end{aligned}
$$

Combining all terms concludes the proof. $\qquad\square$

As an immediate consequence of the above theorem, we have:

Corollary 4.8. *Under the assumptions in section 4.1, the spectrum of the matrix $\mathcal{S}_{\delta t}$, corresponding to the Fourier-transformed forward Euler time-stepper for the kinetic equation (3) (defined in equation (34)) is located in two clusters*

$$
\mathrm{Sp}(S_{\delta t}) \subset \mathcal{D}_1 \cup \mathcal{D}_2
$$

with

$$
\mathcal{D}_1 = \{\tau_{\delta t}\} \ and \ \mathcal{D}_2 = \mathcal{D}\left(1 - \frac{\delta t}{\varepsilon^{\gamma+1}}, \frac{\delta t}{J\varepsilon^\gamma}\max_{j\in\mathcal{J}}(|\alpha_j| + \beta_j|)\right)
$$

The dominant eigenvalue $\tau_{\delta t}$ is simple and can be expanded as

$$
\begin{aligned}
\tau_{\delta t} &= \left(1 - \frac{\delta t}{\varepsilon^{\gamma+1}}\right) + \frac{\delta t}{\varepsilon^{\gamma+1}}\left(1 + \varepsilon\langle\alpha\rangle + \varepsilon^2\left(\langle(\langle\alpha\rangle - \alpha)^2\rangle - \langle\beta^2\rangle + \delta_\gamma\left\langle\frac{\beta}{v}\right\rangle^2\right)\right) \\
&+ \imath\,\frac{\delta t}{\varepsilon^{\gamma+1}}\left(\varepsilon\delta_\gamma\left\langle\frac{\beta}{v}\right\rangle + \varepsilon^2\left(\delta_\gamma\left\langle\left(\left\langle\frac{\beta}{v}\right\rangle - \beta\right)(\langle\alpha\rangle - \alpha)\right\rangle + \delta_{\gamma-1}\left\langle\frac{\beta}{v}\right\rangle\right)\right) + o(\varepsilon^{1-\gamma}).
\end{aligned}
$$

These spectra are illustrated in figure 4, where we have plotted the spectra of the amplification factor of the time-stepper $\mathcal{S}_{\delta t}$ in the spatial domain (see equation (17)) for several choices of δt and for both the parabolic and the hyperbolic scaling.

4.3 Parameter choices for projective integration

Based on the expressions for the spectrum of the inner time-stepper (17) in corollary 4.8 and the stability regions of the projective Runge–Kutta methods in theorem 3.2, we can determine parameter values δt, Δt and K for which the projective Runge–Kutta methods are stable. We first observe from corollary 4.8 that we need

$$
\delta t = \varepsilon^{\gamma+1}
$$

to center the fast eigenvalues of the inner time-stepper (corresponding to the region $\mathcal{D}_2$) around the origin to contain them in the stability region $\mathcal{R}_2$ (see theorem 3.2).
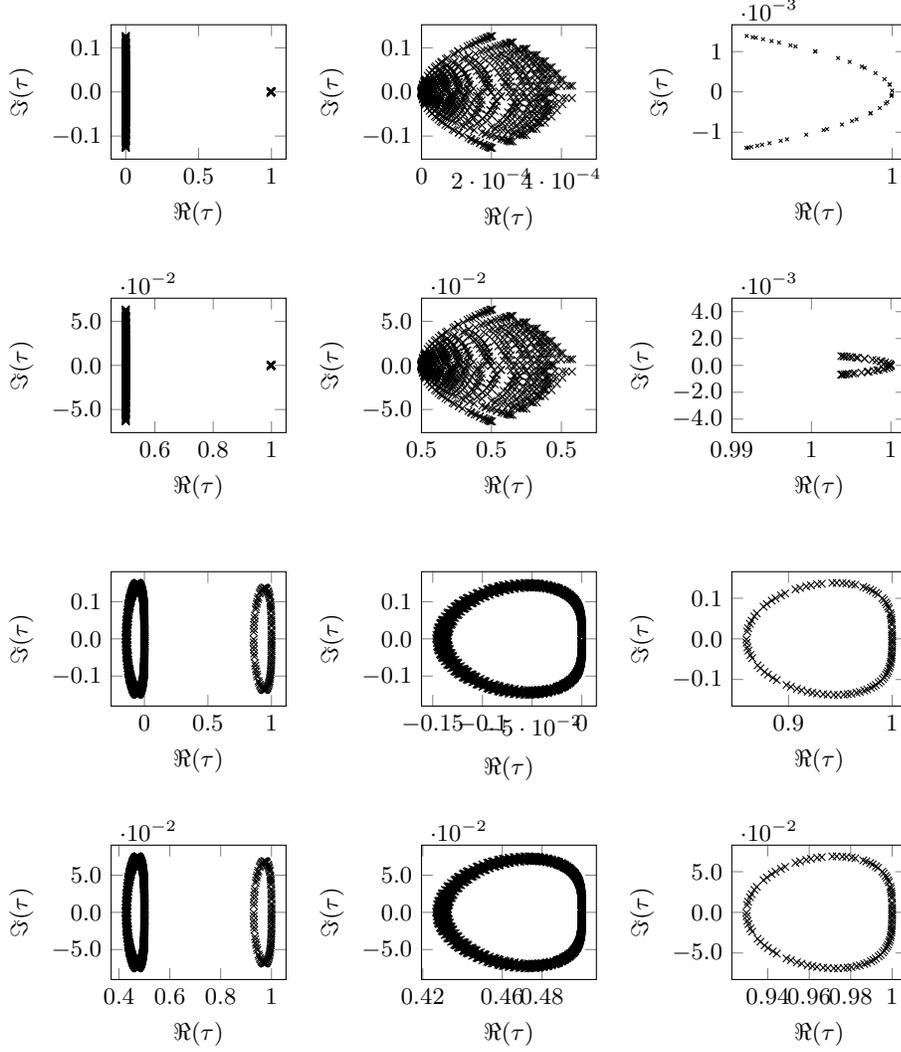
Figure 4: Spectrum of the time-stepper $\mathcal{S}_{\delta t}$. The first two rows correspond to a parabolic scaling ($\gamma = 1$); the last two rows to a hyperbolic scaling ($\gamma = 0$). On each row: Left: a global view of the spectrum; Middle and right: zoom to each of the eigenvalue clusters around 0 and 1. Parameter values are $K = 3, \Delta t = 1 \times 10^{-2}, \Delta x = 0.1, J = 20$ and $\delta t = \varepsilon^{\gamma+1}$ (first row) and $\delta t = 5 \times 10^{-1}\varepsilon^{\gamma+1}$ (second row).

17

Remark 1 (Spatial mesh width). *As observed in [37], we remark that this choice induces a restriction on the spatial mesh width to ensure stability of the inner integrator. Specifically, we require*

$$\frac{\varepsilon}{J} \max_j \left(|\alpha_j| + |\beta_j| \right) \leq 1,$$

from which, using (31) or (32), it follows that $\Delta x \geq C v_J \varepsilon$. However, since we consider the limit when ε tends to 0 for fixed Δx, as we are interested in Asymptotic Preserving schemes, this is not a problematic restriction.

Next, we have to determine Δt such that the slow eigenvalues are captured in $\mathcal{R}_1$ and choose K in such a way that the stability region $\mathcal{R}_2$ is large enough to contain all fast eigenvalues. We have the following conditions:

Theorem 4.9 (Stability of projective Runge–Kutta methods). *When using an inner integrator (17) for the kinetic equation (3) with time step $\delta t = \varepsilon^{\gamma+1}$, a projective Runge–Kutta method (18)-(20) is stable if the macroscopic time step Δt satisfies*

$$\Delta t \leq \varepsilon^\gamma b, \tag{45}$$

and the integer K that determines the number $K + 1$ of inner steps satisfies

$$K \geq \frac{1}{1 + \dfrac{\log(\max_j(|\alpha_j| + |\beta_j|)/J)}{\log(\varepsilon)}} - \frac{\log(b)}{\log(\varepsilon) + \log(\max_j(|\alpha_j| + |\beta_j|)/J)}, \tag{46}$$

$$b = \min\left(\frac{2}{-\langle\alpha\rangle - \varepsilon(\langle((\langle\alpha\rangle - \alpha)^2 - \beta^2\rangle - \delta_\gamma\langle\frac{\beta}{v}\rangle^2)}, \frac{1}{\left| \delta_\gamma\langle\frac{\beta}{v}\rangle + \varepsilon\left[\delta_\gamma\left(\langle\frac{\beta}{v}\rangle - \beta\right)(\langle\alpha\rangle - \alpha) + \delta_{\gamma-1}\langle\frac{\beta}{v}\rangle\right]\right|} \right). \tag{47}$$

Before proceeding to the proof, we make a few observations on the macroscopic time step Δt. At first, consider the hyperbolic scaling ($\gamma = 0$). In this regime, the macroscopic time step Δt is seen to be independent of ε when ε tends to 0. Moreover, since the coefficients α and β depend on $1/\Delta x$, the inequality in condition (45) will result in a CFL-type condition of the form $\Delta t \leq C\Delta x$. Now consider the parabolic scaling ($\gamma = 1$). In that case, the first term in equation (45) can only be bounded independently of ε if $\langle\alpha\rangle = 0$, i.e., by a central scheme. (This is consistent with the observation in [37].) The second term is bounded independently of ε because $\delta_1 = 0$. We then end up with a CFL-type condition of the form $\Delta t \leq C\Delta x^2$. Concrete results for specific schemes are given after the proof. Similarly, the number K of inner steps can be bounded independently of ε using the fact that $\log(\varepsilon) \to -\infty$ as ε tends to 0.

Proof of theorem 4.9. We know from theorem 3.2 that the stability regions of the projective forward Euler method are contained within those of the higher-order Runge–Kutta methods. We therefore can safely choose the method parameters based on the stability conditions for the projective forward Euler method, which are given in equation (24). The chosen method parameters δt, Δt and K need to be chosen to ensure that the eigenvalues in the region $\mathcal{D}_2$ (see corollary 4.8) are contained in the region $\mathcal{D}_2^{PFE}$, and that the eigenvalue $\tau_{\delta t}$ is contained in the region $\mathcal{D}_1$.

First, we center the region $\mathcal{D}_2$ around the origin, resulting in the requirement that

$$\delta t = \varepsilon^{\gamma+1}. \tag{48}$$

Next, we need conditions on Δt such that $\tau_{\delta t}$ is contained within $\mathcal{D}_1^{PFE}$, i.e.

$$1 - 2\frac{\varepsilon^{\gamma+1}}{\Delta t} \leq \text{Re}(\tau_{\delta t}) \leq 1, \qquad |\text{Im}(\tau_{\delta t})| \leq \frac{\varepsilon^{\gamma+1}}{\Delta t},$$

where we have already used (48). The second inequality on $\text{Re}(\tau_{\delta t})$ is always satisfied. Using the expressions for the eigenvalues in corollary 4.8, we obtain

$$
\begin{cases}
1 - 2\dfrac{\delta t}{\Delta t} \leq 1 + \varepsilon\alpha + \varepsilon^2 \left(\left\langle (\langle\alpha\rangle - \alpha)^2 - \beta^2 \right\rangle + \delta_\gamma \left\langle \dfrac{\beta}{v} \right\rangle^2 \right) \\[4mm]
\dfrac{\delta t}{\Delta t} \geq \left| \varepsilon\delta_\gamma \left\langle \dfrac{\beta}{v} \right\rangle + \varepsilon^2 \left[\delta_\gamma \left\langle \left(\left\langle \dfrac{\beta}{v} \right\rangle - \beta \right) (\langle\alpha\rangle - \alpha) \right\rangle + \delta_{\gamma-1} \left\langle \dfrac{\beta}{v} \right\rangle \right] \right|
\end{cases}
$$

from which the condition in (45) is readily satisfied.

Finally, we have to choose K, the number of small steps for the inner integrator, such that the eigenvalues in the region $\mathcal{D}_2$ are contained in the region $\mathcal{D}_2^{PFE}$. From corollary 4.8, we already know that, when $\delta t = \varepsilon^{\gamma+1}$, the radius r of $\mathcal{D}_2$ is given as

$$
r = \frac{\varepsilon}{J} \max_{j \in \mathcal{J}} (|\alpha_j| + \beta_j).
$$

Given that δt and Δt have already been fixed, the stability condition

$$
r \leq \left(\frac{\delta t}{\Delta t} \right)^{1/K} = \left(\frac{\varepsilon^{\gamma+1}}{\Delta t} \right)^{1/K}
$$

results in a condition on K, which can be derived as $K \geq \log\left(\varepsilon^{\gamma+1}/\Delta t \right) / \log(r)$.

Using the conditions we have derived on δt and Δt, we get:

$$
K \geq \frac{\log\left(\varepsilon/b \right)}{\log\left(\varepsilon \, \max_j(|\alpha_j| + |\beta_j|)/J \right)}, \tag{50}
$$

where b is defined as in equation (47). Remarking that (50) is equivalent to (46) concludes the proof. $\qquad\square$

We conclude with the application of the above stability conditions for the specific combinations for the scaling and the spatial discretisation given in (14) and in (15).

Example 1 (Hyperbolic scaling with third order upwind discretisation)**.** *The hyperbolic case corresponds to $\gamma = 0$, which implies $\delta t = \varepsilon$. Given the definitions (31) of α_j and β_j, we obtain the following condition on Δt,*

$$
\begin{aligned}
\Delta t &\leq \min(\Delta t_1^{max}, \Delta t_2^{max}) \\[3mm]
\Delta t_1^{max} &= \left(\frac{2}{\left(\dfrac{\langle |v| \rangle A(\zeta)}{6\Delta x} \right) + \dfrac{\varepsilon}{36\Delta x^2} \left(1 + \langle v^2 \rangle \right) B(\zeta)^2 + \text{Var}(|v|)^2 A(\zeta)^2} \right) \\[3mm]
\Delta t_2^{max} &= \frac{1}{\dfrac{8\sin(\zeta) - \sin(2\zeta)}{6\Delta x} + \varepsilon \left(\langle \alpha\beta \rangle - \langle\alpha\rangle\langle\beta\rangle \right)}
\end{aligned}
$$

It is clear that the order $O(\varepsilon)$ term is positive in the denominator of $\Delta t_1^{\max}$, and 0 in the denominator of $\Delta t_2^{\max}$, since $\langle \alpha\beta \rangle - \langle\alpha\rangle\langle\beta\rangle = 0$. In the limit when ε tends to 0, we then obtain the following condition on Δt:

$$
\Delta t \leq \min\left(\frac{3\Delta x}{4\langle |v| \rangle}, \frac{3\Delta x}{8} \right).
$$

We end up with a stability condition on the macroscopic time step Δt which is independent of ε, and that is of CFL-type for a hyperbolic partial differential equation.

To bound the number of inner steps K, we observe that, when ε tends to 0, the second term in (46) tends to 0. Moreover, assuming $(1/J)\max_j(|\alpha_j| + |\beta_j|) \leq 1$, the first term is bounded by 1. With some algebraic manipulation, this leads to the condition that $K \geq 2$.

19

Example 2 (Parabolic scaling with fourth order central discretisation)**.** *The parabolic case corresponds to* $\gamma = 1$*, and therefore* $\delta t = \varepsilon^2$*. For the fourth order central discretisation,* α_j *and* β_j *are given by* (32).

Substituting these expressions into (45) *yields*

$$
\begin{aligned}
\Delta t &\leq& \min(\Delta t_1^{max}, \Delta t_2^{max}) \\
\Delta t_1^{max} &=& \frac{2\varepsilon(6\Delta x)^2}{\varepsilon(8\sin(\zeta) - \sin(2\zeta))^2} \geq \frac{9\Delta x^2}{8\langle v^2 \rangle} \\
\Delta t_2^{max} &=& \frac{\varepsilon 6\Delta x}{\varepsilon(8\sin(\zeta) - \sin(2\zeta))} \geq \frac{3\Delta x}{4}
\end{aligned}
$$

Concerning the number K *of inner steps, a similar argument as above can be followed to show that the projective Runge–Kutta scheme will be stable provided that* $K \geq 3$*, see also [37].*

5   Consistency analysis

In this section we will prove that the PRK4 algorithm is fourth order accurate in space and time for a linear flux $A(u) = u$. [1] First let us introduce some notations that will be used throughout this section: for $k \in \{0, \ldots, K+1\}$,

- $t^{N,k} = N\Delta t + k\delta t$ is an intermediate time on the micro grid, as described in subsection 3.2,

- $\widetilde{\partial^p f}^{N,k}$ denotes the evaluation of a p-th derivative of the exact solution of (2) at time $t^{N,k}$,

- and $\tilde{u}^{N,k} = \langle \tilde{f}^{N,k} \rangle$ is the corresponding exact density,

- while $f^{N,k}$ is the numerical solution at time $t^{N,k}$ resulting from the PRK4 scheme, starting from the exact solution $\tilde{f}^{N,K}$

- Similarly $u^{N,k} = \langle f^{N,k} \rangle$ is the corresponding numerical density function.

Therefore we will compute the truncation error E^{N+1} at time t^{N+1} which is defined as:

$$
E^{N+1} = \frac{\tilde{u}^{N+1} - u^{N+1}}{\Delta t}, \tag{51}
$$

The expression for the truncation error the PRK4 scheme is:

$$
E^{N+1} = \frac{\tilde{u}^{N+1} - u^{N,K+1}}{\Delta t} - \frac{\Delta t - (K+1)\delta t}{\Delta t} \sum_{s=1}^{S} b_s \frac{u^{N+c_s,K+1} - u^{N+c_s,K}}{\delta t},
$$

with, $\forall s \in \{1, \ldots, S\}$,

$$
u^{N+c_s} = u^{N,K+1} + (c_s\Delta t - (K+1)\delta t) \sum_{l=1}^{s-1} \frac{a_{sl}}{c_s} k_l,
$$

where, $\forall l \in \{1, \ldots, s-1\}$,

$$
k_l = \frac{u^{N+c_l,K+1} - u^{N+c_l,K}}{\delta t}.
$$

Furthermore, the convergence error for the inner integrator reads:

$$
e_f^{N,k} := \frac{\tilde{f}^{N,k} - f^{N,k}}{\delta t}. \tag{52}
$$

Recall that, since $\delta t = \varepsilon^{\gamma+1}$,

$$
f^{N,k+1} = \mathcal{S}_{\delta t} f^{N,k} = -\varepsilon \mathcal{D}_{x,v}(f^{N,k}) + \mathcal{M}_v u^{N,k}. \tag{53}
$$

---

[1]Remark that the following analysis can be done for PRK schemes of any order.

Remark 2. *To stress the fact that $\mathcal{M}_v$ is a linear operator, we omit the parenthesis of the argument in this section.*

Now we want to analyse the evolution of the truncation error of the inner integrator:

Lemma 5.1. *Suppose, we use an inner integrator which is accurate up to p-th order in space and first order in time. Then, we also have that $e^{N,K} = O(\delta t) + O(\varepsilon^{1-\gamma}\Delta x^p)$.*

Proof. First, we analyse how the truncation error, defined in (52) evolves after one extra step with the inner integrator. Furthermore, we can expand the exact solution $\tilde{f}^{N,k+1}$ at time $t^{N,k+1}$ around $t^{N,k}$ by using Taylor series:

$$
\begin{aligned}
\tilde{f}^{N,k+1} &= \tilde{f}^{N,k} + \delta t\,\widetilde{\partial_t f}^{N,k} + O(\delta t^2) \\
&= \tilde{f}^{N,k} + \frac{\delta t}{\varepsilon^\gamma}\left(-v\widetilde{\partial_x f}^{N,k} + \frac{-\tilde{f}^{N,k} + \mathcal{M}_v\langle \tilde{f}^{N,k}\rangle}{\varepsilon}\right) + O(\delta t^2) \\
&= \mathcal{S}_{\delta t}\tilde{f}^{N,k} + \frac{1}{\varepsilon^\gamma}(\mathcal{D}_{x,v}\tilde{f}^{N,k} - v\widetilde{\partial_x f}^{N,k}) + O(\delta t^2)
\end{aligned}
\tag{54}
$$

Using (53) and (54), we get

$$
e^{N,k+1} = \mathcal{S}_{\delta t}e^{N,k} + \frac{1}{\varepsilon^\gamma}\left(\mathcal{D}_{x,v}(\tilde{f}^{N,k}) - v\widetilde{\partial_x f^{N,k}}\right) + O(\delta t).
$$

Recall that we suppose that the inner integrator is stable, and the assumption that the result at time t^N is exact. This implies that we can write $e^{N,K+1}$ as:

$$
e^{N,K+1} = \frac{1}{\varepsilon^\gamma}\sum_{k=0}^{K-k}\mathcal{S}_{\delta t}^k(\mathcal{D}_{x,v}\tilde{f}^{N,k} - v\widetilde{f}^{N,k}) + O((K+1)\delta t).
\tag{55}
$$

To consider the above expression in more detail, we define: $\Delta : f \mapsto \mathcal{D}_{x,v}(\tilde{f}) - \widetilde{v\partial_x f}$ and recall that $\mathcal{D}_{x,v}$ and $\mathcal{M}_v$ are linear operators. Since

$$
\mathcal{S}_{\delta t}(\Delta f) = \mathcal{M}_v\langle\Delta f\rangle - \varepsilon\mathcal{D}_{x,v}(\Delta f),
$$

a simple recursion leads to, for all $k \geq 2$,

$$
\mathcal{S}_{\delta t}^k(\Delta f) = \mathcal{M}_v\langle\Delta f\rangle - \varepsilon\{\mathcal{D}_{x,v}\mathcal{M}_v\langle\Delta f\rangle + \mathcal{M}_v\langle\Delta f\rangle\} + O(\varepsilon^2).
$$

Now taking the mean value over velocity space yields:

$$
\langle\Delta f\rangle = \langle\mathcal{D}_{x,v}(\tilde{u} + \varepsilon\tilde{g}) - v\widetilde{\partial_x(u + \varepsilon g)}\rangle = \varepsilon\langle\Delta g\rangle = \varepsilon O(\Delta x^p).
$$

The proof of the statement then follows by a simple substitution of the above estimate into equation (55). $\square$

Now we can finally calculate the truncation error.

Theorem 5.2 (Truncation Error of PRK scheme). *Consider a PRK scheme, that satisfies the assumptions 3.1 and (25) on the coefficients $a_{s,l}$. Then the truncation error E^{N+1} of the scheme can be described by:*

$$
\begin{aligned}
E^{N+1} &= \left(\Delta t^2\mu^3\left(\frac{1}{6} - b^T A^2 e\right) + \Delta t^3\mu^4\left(\frac{1}{24} - b^T A^3 e\right)\right)f^N + O(\Delta t^4) \\
&+ O(\delta t) + O(\varepsilon^{1-\gamma}\Delta x^p) + \delta t O\left(\frac{\delta t + \varepsilon^{1-\gamma}\Delta x^p}{\Delta t}\right),
\end{aligned}
$$

where μ indicates the eigenvalue of the inner integrator.

21

Proof. First we will derive a relation between the derivatives $k_s = \partial_t(f^{N+c_s})$ from the original Runge-Kutta scheme and the modified derivatives for our PRK scheme. So let us perform a Taylor expansion from k_s:

$$k_s = \frac{f^{N+c_s} + (K+1)\delta t\mathcal{D}_t(f^{N+c_s}) + (K+1)^2\delta t^2\mathcal{D}_{tt}(f^{N+c_s})/2 - f^{N+c_s} - K\delta t\mathcal{D}_t(f^{N+c_s}) - (K^2\delta t^2/2)\mathcal{D}_{tt}(f^{N+c_s})}{\delta t}$$

$$= \mathcal{D}_t(f^{N+c_s}) + \frac{2k+1}{2}\delta t\mathcal{D}_{tt}(f^{N+c_s}) + O(\delta t^2).$$

Now we showed in lemma 5.1 that an application of the numerical derivative introduces an error of order $O(\varepsilon^{1-\gamma}\Delta x^p)$ with respect to the exact derivative ∂_t and hence we can write:

$$k_s = \partial_t(f^{N+c_s}) + \frac{2K+1}{2}\delta t\partial_{tt}(f^N + c_s) + O(\varepsilon^{1-\gamma}\Delta x^p).$$

We proceed by substituting the expression for f^{N+c_s} into the above equation. This yields:

$$k_s = \partial_t\left(f^{N,K+1} + (c_s\Delta t - (K+1)\delta t)\sum_{l=1}^{s-1}\frac{a_{sl}}{c_s}k_l\right) + O(\delta t + \varepsilon^{1-\gamma}\Delta x^p).$$

Of course, the latter can be further expanded as follows:

$$k_s = \partial_t\left(f^N + (K+1)\delta t\mathcal{D}_t(f^N) + (c_s\Delta t - (K+1)\delta t)\sum_{l=1}^{s-1}\frac{a_{sl}}{c_s}k_l\right) + O(\delta t + \varepsilon^{1-\gamma}\Delta x^p),$$

which is in turn equivalent to:

$$k_s = \partial_t(f^N) + (K+1)\delta t\left(\partial_{tt}(f^N) - \sum_{l=1}\frac{a_{sl}}{c_s}k_l\right) + \Delta t\sum_{l=1}^{s-1}a_{sl}k_l + (1+\delta t)O(\delta t + \varepsilon^{1-\gamma}\Delta x^p).$$

Next, a combination of this result with the equation for f^{N+1} gives rise to:

$$\begin{aligned}
f^{N+1} &= f^N + (K+1)\delta t\mathcal{D}_t f^N + \frac{(K+1)^2}{2}\delta t^2\mathcal{D}_{tt}(f^N) \\
&+ (\Delta t - (K+1)\delta t)\sum_{s=1}b_s\left(\partial_t(f^N) + (K+1)\delta t\left(\partial_{tt}(f^N) - \sum_{l=1}\frac{a_{sl}}{c_s}k_l\right) + \Delta t\sum_{l=1}^{s-1}a_{sl}k_l + (1+\delta t)O(\delta t + \varepsilon^{1-\gamma}\Delta x^p)\right).
\end{aligned}$$

Then, we will proceed by splitting the above expression in an δt -independent part and a part which depends on the time step of the inner integrator. Now, we can apply the theorem about the order conditions of general RK schemes (see [26]) to derive finally the expression for the truncation error of the PRK scheme. $\qquad\square$

Example 3 (Truncation error for PRK4). *As a direct consequence of theorem 5.2, we find that the order of accuracy of the PRK4 scheme is:*

$$E^{N+1} = O(\Delta t^4) + O(\delta t) + O(\varepsilon^{1-\gamma}\Delta x^p) + \delta t\left(\frac{\delta t + \varepsilon^{1-\gamma}\Delta x^p}{\Delta t}\right),$$

where we have used that $a_{21} = a_{32} = 1/2$ and $a_{43} = 1$. The other Runge Kutta coefficients are zero. Moreover the scheme is consistent. This is the scheme that we used throughout the numerical experiments.

6 Numerical experiments

In this section, we illustrate the performance of the high-order projective integration algorithm. In section 6.1, we first illustrate its consistency properties and long term performance on a simple linear kinetic equation. Afterwards, we will apply the scheme on some more realistic applications: the Burgers' equation (section 6.2) and the semiconductor equation (section 6.3).

In sections 6.1 and 6.2, we consider the velocity space $v_j, j \in \{1 \cdots J\}$ to be constructed using the zeroes of the Legendre polynomial of degree J; in section 6.3, we use the zeroes of the Hermite polynomial of degree $J = 20$. All simulations are performed on the spatial domain $[-1, 1]$. We choose an equidistant, constant in time mesh with cell centers $\Pi := \{x_0 = -1 + \Delta x/2, \ldots, 1 - \Delta x/2\}$, and the fourth-order central spatial discretisation defined by (14).

6.1 Linear kinetic equation

We consider equation (3) with $A(u) = u$ and periodic boundary conditions. As an initial condition, we take

$$f(v, x, t) = \frac{e^{-v^2 \sin(\pi x)/T}}{\sum_j w_j}, \qquad j = 1, \ldots, J.$$

To examine the truncation error (defined in equation (51)), we perform a numerical simulation using a second and fourth order PRK algorithm, with Butcher tableaux in figure 2(right), using $\delta t = \varepsilon^2$, $K = 3$, and $\Delta t = 1 \times 10^{-3}$, and $\Delta x = 1 \times 10^{-2}$; the number of outer PRK steps is defined by $(N + 1)\Delta t = 1$. We perform the experiment for $\varepsilon = 1 \times 10^{-2}$ and $\varepsilon = 1 \cdot 10^{-3}$. As the reference solution, we use a direct forward Euler simulation with $\delta t = \varepsilon^3$. The results are shown in figure 5. One observes that the truncation error behaves as $O(\Delta t^4)$, resp., $O(\Delta t^2)$, for large Δt until, for sufficiently small Δt, a plateau is reached, at which the contribution of the inner integrator to the truncation error, which is $O(\varepsilon^2)$, becomes dominant.
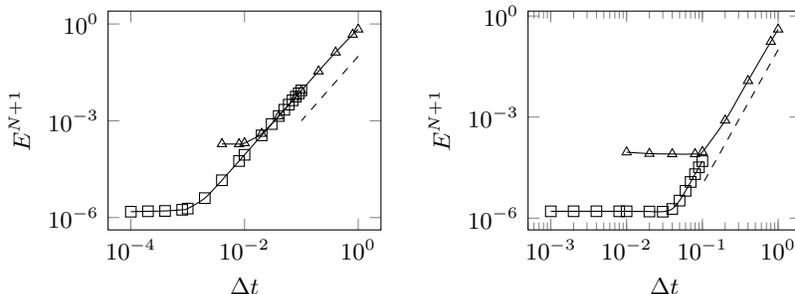

Figure 5: Truncation error of u with PRK4 as a function of Δt for $K = 3, \Delta x = 1 \times 10^{-2}$ using $\varepsilon = 1 \times 10^{-3}$ (squares) and $\varepsilon = 1 \times 10^{-2}$ (triangles).

Next, we compare the long-time simulation results of the PRK scheme with both a full microscopic simulation and a simulation of the limiting macroscopic equation (11). We consider $\varepsilon = 1 \times 10^{-3}$, and choose a fourth-order PRK scheme with $K = 3$, $\delta t = \varepsilon^2$, $\Delta x = 1 \times 10^{-1}$ and $\Delta t = 1 \times 10^{-3}$. The full microscopic simulation is performed using the same inner integrator with time step $\delta t = \varepsilon^3$, whereas the limiting macroscopic equation is simulated using the same spatial discretization and the corresponding direct Runge–Kutta method of order 4. The results are shown in figure 6. We observe that the PRK4 algorithm is visually as accurate as the full microscopic simulation, while requiring a computational effort that is only 1/1000 of the full microscopic simulation. Moreover, the projective scheme appears to be able to capture the kinetic behaviour that is lost in the macroscopic limiting equation.
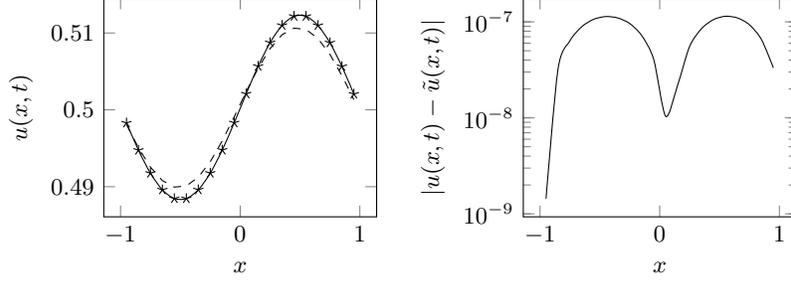
Figure 6: Long term results for the linear kinetic equations. Left: density at time 1, $\Delta t = 1 \times 10^{-3}$, $K = 3$, $\Delta x = 1 \times 10^{-1}$ with PRK4 (stars); microscopic evolution with $\delta t = \varepsilon^3$ (solid line); results obtaind using the limiting equation (dashed). Right: absolute error of PRK results with respect to full microscopic simulation.
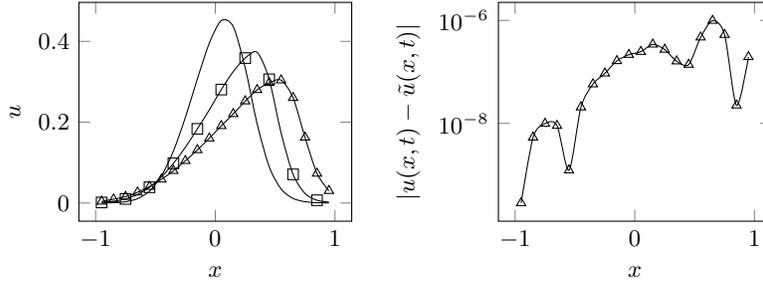

Figure 7: Left: Evolution of the Viscous Burgers equation as a result of the PRK4 algorithm for $N = 100$ (solid line, no marks), $N = 500$ (squares) and $N = 1000$ (triangles). Right: Error at $N = 1000$. Parameters: $\varepsilon = 1 \times 10^{-3}, \Delta t = 1 \times 10^{-3}, K = 3, \Delta x = 1 \times 10^{-1}$.

6.2 Viscous Burgers' equation

Let us now consider the viscous Burgers' equation, i.e., equation (3) with $A(u) = u^2$, using Neumann boundary conditions

$$\partial_x f(-1, v, t) = \partial_x f(1, v, t) = 0),$$

and the initial condition:

$$f(x, v_j, t) = \frac{1}{\sum_j w_j} \exp(-v_j^2/T) \exp(-x^2/0.1). \tag{57}$$

We again perform a fourth order PRK simulation using $K = 3$, $\delta t = \varepsilon^2$, $\Delta x = 1 \times 10^{-1}$ and $\Delta t = 1 \times 10^{-3}$. As a reference solution, we perform a full microscopic simulation using the same inner integrator with time step $\delta t = \varepsilon^3$. The results are shown in figure 7(left) at various instances in time. On the right, the error with respect to the reference soluton is shown. We clearly observe that the projective integration method is also very accurate in this case.

6.3 Semiconductor equation

Finally, we illustrate the PRK method for the semiconductor equation (12) with $\varepsilon = 1 \times 10^{-3}$. To discretise the partial derivative ∂_v, we also use a second order finite difference scheme, taking into account that the chosen velocities, because they are the zeroes of the Hermite polynomials, are not equidistant. As an initial condition, we again choose (57) and we apply no-flux boundary conditions for both the velocity and spatial variables. For the potential Φ, we applied Dirichlet boundary conditions, $\Phi(-1, t) = 2.0$ and $\Phi(1, t) = 0$, causing an advective movement to the left.
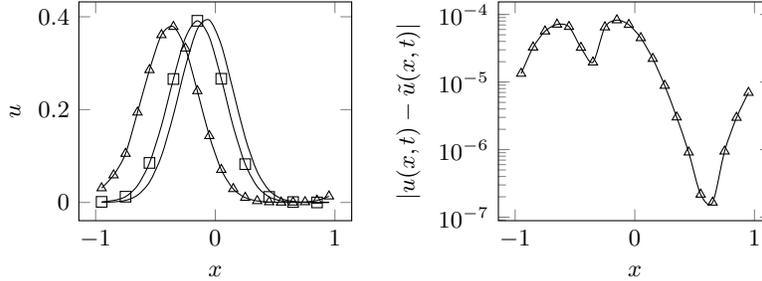
24

Figure 8: Evolution of the density calculated with the projective integration algorithm after $N = 100$ (no marks), $N = 200$ (squares), $N = 500$ (triangles). Parameters of the simulation: $\varepsilon = 1 \times 10^{-3}$, $\Delta t = 1 \times 10^{-3}$, $\Delta x = 1 \times 10^{-1}$, $\Phi(-1, t) = -2.0$, $\Phi(1, t) = 0.0$, $T = 1 \times 10^{-2}$.

As before, we perform a fourth order PRK simulation using $K = 3$, $\delta t = \varepsilon^2$, $\Delta x = 1 \times 10^{-1}$ and $\Delta t = 1 \times 10^{-3}$, as well as a microscopic reference solution using the same inner integrator with time step $\delta t = \varepsilon^3$. The results are shown in figure 8(left) at various instances in time. On the right, the error with respect to the reference soluton is shown. We clearly observe that the projective integration method is also very accurate in this case.

7 Conclusions

We investigated a high-order, fully explicit, asymptotic-preser-ving scheme for a kinetic equation with linear relaxation, both in the hydrodynamic and diffusive scalings in which a hyperbolic, resp. parabolic, limiting equation exists. The scheme first takes a few small (inner) steps with a simple, explicit method (such a direct forward Euler) to damp out stiff components of the solution and estimate the time derivative of the slow components. These estimated time derivatives are then used in an (outer) Runge–Kutta method of arbitrary order. We showed that, with an appropriate choice of inner step size, the time-step restriction on the outer time step is similar to the stability condition for the limiting macroscopic equations. Moreover, the number of inner time steps is also independent of the scaling parameter. We analyzed stability and consistency, and illustrated with numerical results.

We conclude by pointing out the current limitations of the method, and some suggestions for future work. The asymptotic-preserving nature of the scheme is due to the presence of a single relaxation time in the linear relaxation collision operator, and relies on an appropriate choice of the inner time step, which has to satisfy $\delta t = \varepsilon^{\gamma+1}$. When multiple relaxation times are present, one should expect the number of time steps to be chosen as $K \sim \log(1/\varepsilon)$ [18]. In such situations, it might be of interest to study schemes in which a sequence of inner steps is taken, each commensurate with one of the relaxation time scales, as is proposed in [16]. A second direction of further investigation would be to look at problems in which hydrodynamic and diffusive regimes are present simultaneously in different parts of the spatial domain. Many efforts have been done for (semi-)implicit asymptotic-preserving schemes (some of them cited in the introduction); a number of these techniques (such as an a priori modeling of boundary layers [21, 53]) can be readily applied in conjunction with the projective integration method proposed here.

References

[1] Marvin Adams. Discontinuous Finite Element Transport Solutions in Thick Diffusive Problems. *Journal Name: Nuclear Science and Engineering*, 2001.

[2] D Aregba-Driollet and R Natalini. Discrete Kinetic Schemes for Multidimensional Systems of Conservation Laws. *SIAM J. Numer. Anal.*, 37(6):1973–2004, 2000.

[3] Uri M Ascher, Steven J Ruuth, and Raymond J Spiteri. Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations. *Applied Numerical Mathematics*, 25(2–3):151–167, 1997.

[4] Christophe Besse and Thierry Goudon. Derivation of a Non-Local Model for Diffusion Asymptotics—Application to Radiative Transfer Problems. *Commun. Comput. Phys.*, 8(5):1139, 2010.

[5] P L Bhatnagar, E P Gross, and M Krook. A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems. *Physical Review*, 94(3):511–525, 1954.

[6] S Boscarino, L Pareschi, and G Russo. Implicit-Explicit Runge–Kutta Schemes for Hyperbolic Systems and Kinetic Equations in the Diffusion Limit. *SIAM J. Sci. Comput.*, 35(1):A22–A51, 2013.

[7] F Bouchut. Construction of BGK Models with a Family of Kinetic Entropies for a Given System of Conservation Laws. *J. Stat. Phys.*, 95(1):113–170, 1999.

[8] Christophe Buet and Stéphane Cordier. An asymptotic preserving scheme for hydrodynamics radiative transfer models. *Numer. Math.*, 108(2):199–221, 2007.

[9] Christophe Buet and Bruno Despres. Asymptotic preserving and positive schemes for radiation hydrodynamics. *J. Comput. Phys.*, 215(2):717–740, 2006.

[10] J. a. Carrillo, T. Goudon, P. Lafitte, and F. Vecil. Numerical Schemes of Diffusion Asymptotics and Moment Closures for Kinetic Equations. *J. Sci. Comput.*, 36(1):113–149, 2008.

[11] H Cohen. *Numerical Approximation Methods*. Springer, 2011.

[12] J.-F. Coulombel, F Golse, and T Goudon. Diffusion approximation and entropy-based moment closure for kinetic equations. *Asymptot. Anal.*, 45:1–39, 2005.

[13] G Dimarco and L Pareschi. Asymptotic Preserving Implicit-Explicit Runge–Kutta Methods for Nonlinear Kinetic Equations. *SIAM J. Numer. Anal.*, 51(2):1064–1087, 2013.

[14] W E and B Engquist. The heterogeneous multi-scale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[15] W E, B Engquist, X Li, W Ren, and E Vanden-Eijnden. Heterogeneous multiscale methods: A review. *Commun. Comput. Phys.*, 2(3):367–450, 2007.

[16] K Eriksson, C Johnson, and A Logg. Explicit Time-Stepping for Stiff ODEs. *SIAM J. Sci. Comput.*, 25(4):1142–1157, 2004.

[17] Francis Filbet and Shi Jin. An Asymptotic Preserving Scheme for the ES-BGK Model of the Boltzmann Equation. *J. Sci. Comput.*, 46(2):204–224, 2010.

[18] C. W. Gear and Ioannis G Kevrekidis. Projective Methods for Stiff Differential Equations: Problems with Gaps in Their Eigenvalue Spectrum. *SIAM J. Sci. Comput.*, 24(4):1091–1106, 2003.

[19] Alì Giuseppe and Angelo Marcello Anile. Moment equations for charged particles: global existence results. In Pierre Degond, Lorenzo Pareschi, and Giovanni Russo, editors, *Modeling and Computational Methods for Kinetic Equations*, Modeling and Simulation in Science, Engineering and Technology, pages 59–80. Birkhäuser Boston, 2004.

[20] P Godillon-Lafitte and T Goudon. A Coupled Model for Radiative Transfer: Doppler Effects, Equilibrium, and Nonequilibrium Diffusion Asymptotics. *Multiscale Modeling \& Simulation*, 4(4):1245–1279, 2005.

[21] F Golse, S Jin, and C Levermore. The Convergence of Numerical Transfer Schemes in Diffusive Regimes I: Discrete-Ordinate Method. *SIAM J. Numer. Anal.*, 36(5):1333–1369, 1999.

[22] Laurent Gosse and G Toscani. Space Localization and Well-Balanced Schemes for Discrete Kinetic Models in Diffusive Regimes. *SIAM J. Numer. Anal.*, 41(2):641–658, 2003.

[23] Laurent Gosse and Giuseppe Toscani. Asymptotic-preserving \& well-balanced schemes for radiative transfer and the Rosseland approximation. *Numer. Math.*, pages 223–250, 2004.

[24] J Guermond and G Kanschat. Asymptotic Analysis of Upwind Discontinuous Galerkin Approximation of the Radiative Transport Equation in the Diffusive Limit. *SIAM J. Numer. Anal.*, 48(1):53–78, 2010.

[25] Jeffrey R Haack and Cory D Hauck. Oscillatory behavior of asymptotic-preserving splitting methods for a linear model of diffusive relaxation. *Kinet. Relat. Models*, 1(4):573–590, 2008.

[26] Ernst Hairer, Gerhard Wanner, and Syvert Norsett. *Solving Ordinary Differential Equations I*. Springer Berlin Heidelberg, 1993.

[27] C Hauck and R Lowrie. Temporal Regularization of the P_N Equations. *Multiscale Modeling & Simulation*, 7(4):1497–1524, 2009.

[28] S Jin, L Pareschi, and G Toscani. Diffusive Relaxation Schemes for Multiscale Discrete-Velocity Kinetic Equations. *SIAM J. Numer. Anal.*, 35(6):2405–2439, 1998.

[29] S Jin, L Pareschi, and G Toscani. Uniformly Accurate Diffusive Relaxation Schemes for Multiscale Transport Equations. *SIAM J. Numer. Anal.*, 38(3):913–936, 2000.

[30] S Jin, Z Xin, Shi Jin, and Zhouping Xin. The Relaxation Schemes for Systems of Conservation Laws in Arbitrary Space Dimensions. *Comm. Pure Appl. Math.*, 48:235–277, 1995.

[31] Shi Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, pages 1–24, 1999.

[32] I G Kevrekidis, C W Gear, J M Hyman, P G Kevrekidis, O Runborg, and C Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level tasks. *Commun. Math. Sci.*, 1(4):715–762, 2003.

[33] Ioannis G Kevrekidis and Giovanni Samaey. Equation-free multiscale computation: algorithms and applications. *Annual review of physical chemistry*, 60:321–44, 2009.

[34] A Klar. A numerical method for kinetic semiconductor equations in the drift-diffusion limit. *SIAM J. Sci. Comput.*, 20(5):1696–1712, 1998.

[35] A Klar. An Asymptotic-Induced Scheme for Nonstationary Transport Equations in the Diffusive Limit. *SIAM J. Numer. Anal.*, 35(3):1073–1094, 1998.

[36] A Klar. An Asymptotic Preserving Numerical Scheme for Kinetic Equations in the Low Mach Number Limit. *SIAM J. Numer. Anal.*, 36(5):1507–1527, 1999.

[37] Pauline Lafitte and Giovanni Samaey. Asymptotic-preserving Projective Integration Schemes for Kinetic Equations in the Diffusion Limit. *SIAM J. Sci. Comput.*, 34(2):A579–A602, 2012.

[38] Edward W Larsen and J.E. Morel. Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes II. *J. Comput. Phys.*, 83(1):212–236, 1989.

[39] Steven L Lee and C William Gear. Second-order accurate projective integrators for multiscale problems. *J. Comput. Appl. Math.*, 201(1):258–274, 2007.

[40] M Lemou and L Mieussens. A New Asymptotic Preserving Scheme Based on Micro-Macro Formulation for Linear Kinetic Equations in the Diffusion Limit. *SIAM J. Sci. Comput.*, 31(1):334–368, 2008.

[41] R B Lowrie and J E Morel. Discontinuous Galerkin for hyperbolic systems with stiff relaxation. In *Discontinuous Galerkin methods ,Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 385–390. Springer, Berlin, 2000.

[42] R G McClarren and R B Lowrie. The effects of slope limiting on asymptotic-preserving numerical methods for hyperbolic conservation laws. *J. Comput. Phys.*, 227(23):9711–9726, 2008.

[43] Ward Melis and Giovanni Samaey. A relaxation method with projective integration for solving nonlinear systems of hyperbolic conservation laws. 2013.

[44] Gerald N Minerbo. Maximum entropy Eddington factors. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 20(6):541–545, 1978.

[45] G C Pomraning. *Linear kinetic theory and particle transport in stochastic mixtures*, volume 7 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co. Inc., River Edge, NJ, 1991.

[46] R Rico-Mart\'inez, C W Gear, and Ioannis G Kevrekidis. Coarse projective kMC integration: forward/reverse initial and boundary value problems. *J. Comput. Phys.*, 196(2):474–489, 2004.

[47] Sergei L Shmakov. A universal method of solving quartic equations. *Int. J. Pure Appl. Math.*, 71(2):251–259, 2011.

[48] B P Sommeijer. Increasing the real stability boundary of explicit methods. *Computers \& Mathematics with Applications*, 19(6):37–49, 1990.

[49] G Strang. On the Construction and Comparison of Difference Schemes. *SIAM J. Numer. Anal.*, 5(3):506–517, 1968.

[50] H Struchtrup. *Macroscopic transport equations for rarefied gas flows: approximation methods in kinetic theory.* Springer, 2005.

[51] David Ulrich. *Complex Made Simple.* American Mathematical Society, 2008.

[52] C Vandekerckhove, D Roose, and K Lust. Numerical stability analysis of an acceleration scheme for step size constrained time integrators. *J. Comput. Appl. Math.*, 200(2):761–777, 2007.

[53] Marie H{\'{e}}lene Vignal. A Boundary Layer Problem for an Asymptotic Preserving Scheme in the Quasi-Neutral Limit for the Euler–Poisson System. *SIAM J. Appl. Math.*, 70(6):1761–1787, 2010.

A Parametrization of stability regions

We need to derive the expressions that are given in Proposition 3.3. Let us start from the projective Runge–Kutta method as applied to the linear test equation, i.e., equation (58), which we now write as

$$
\begin{cases}
k_1 & =: \kappa_1(\tau)y^n = \dfrac{\tau^{K+1} - \tau^K}{\delta t} y^n \\
k_s & =: \kappa_s(\tau)y^n = \dfrac{\tau^{K+1} - \tau^K}{\delta t} F_s(\tau)y^n, \qquad 2 \le s \le S, \\
y^{n+1} & =: \sigma(\tau)y^n = \left(\tau^{K+1} + (M\delta t)\sum_{s=1}^{S} b_s \kappa_s\right) y^n,
\end{cases}
\tag{58}
$$

with $F_s(\tau) = \left(\tau^{K+1} + (M_s\delta t)\sum_{l=1}^{s-1}\dfrac{a_{s,l}}{c_s}\kappa_l\right)$. The stability boundary of the projective Runge–Kutta method is then given by all values of τ such that

$$
\|\sigma(\tau)\| = \left|\tau^{K+1} + (M\delta t)\sum_{s=1}^{S} b_s \kappa_s\right| = 1.
\tag{59}
$$

For small values of $z = \delta t / \Delta t$, the stability region consists of two parts, one part close to the origin, and one part close to 1. To locate these stability boundaries for small z, we proceed in two steps:

(i) We notice that $\sigma(\tau)$ depends on $\kappa_s(\tau)$, while the $\kappa_s(\tau)$ themselves are recursively defined. We therefore first obtain an explicit formula for each of the quantities $\kappa_s(\tau)$, such that we have an explicit formula for $\sigma(\tau)$ as a function of τ.

(ii) Next, for each of the stability regions, we perform an asymptotic expansion of τ as a function of z, and impose the condition (59).

Before proceeding with the derivation, we introduce some additional notation. We will denote by $\bar{A}$ the matrix

$$\bar{A} = \bar{C}A = \operatorname{diag}\left(M_1/c_1, M_2, \ldots, M_s/c_s\right) A$$

where A denotes the matrix of RK-coefficients corresponding to a general S-stage Runge–Kutta scheme. We also introduced the vectors e_1 and e_r which are defined as $(1, 0, \ldots)^T$ and $(0, 1, 1, \ldots)^T$ respectively.

Step (i): Derivation of expression for $F_s(\tau)$ Let us now first derive an expression for $F_s(\tau)$. We will show that

$$\begin{cases} F_s(\tau) = 1 & s = 1, \\ F_s(\tau) = \sum_{j=0}^{s-1}(\tau^K(\tau-1))^j((\bar{A}^j)(e_1 + \tau^{K+1}e_r))_s & \forall 2 \le s \le S. \end{cases} \tag{60}$$

By definition, we know that F_s can be written as:

$$F_s(\tau) = \tau^{K+1} + \sum_{l=1}^{s-1} \bar{a}_{sl} F_l(\tau^{K+1} - \tau^K),$$

where we have introduced $\bar{a}_{sl} = M_s a_{sl}/c_s$ to avoid notational complexity. Then we can derive equation (60) by induction.

Base step ($s = 2$). By definition the following holds:

$$F_2(\tau) = \tau^{K+1} + a_{21}(\tau^{K+1} - \tau^K).$$

Now remark that $(\mathcal{I}e_1)_2 = 0$, while $(\mathcal{I}e_r)_2 = 1$. Then the statement follows from a simple substitution and rearranging the terms.

Induction step. We impose that equation (60) is valid for all $l = 1, \ldots s-1$ as induction hypothesis. So, let's consider $F_s(\tau)$ in more detail:

$$\begin{aligned} F_s(\tau) &= \tau^{K+1} + \sum_{l=1}^{s-1} \bar{a}_{sl} \sum_{j=0}^{l-1} \left(\tau^K(\tau-1)\right)^j \left((\bar{A}^j e_1)_l + (\bar{A}^j e_r)_l \tau^{K+1}\right) \tau^K(\tau-1) \\ &= \tau^{K+1} + \sum_{l=1}^{s-1} \bar{a}_{sl}\tau^K(\tau-1) + \left(\tau^K(\tau-1)\right)^2 \sum_{l=2}^{s-1} \bar{a}_{sl}\left(\bar{A}e_1 + \tau^{K+1}\bar{A}e_r\right)_l + \ldots \\ &= \tau^{K+1} + \sum (\tau^K(\tau-1))^l (\bar{A}^l e_1 + \tau^{K+1}\bar{A}^l e_r)_s \end{aligned}$$

and hence, the equation for the amplification factor σ_S reads:

$$\sigma_S(\theta) = \tau^{K+1} + M\sum_{s=1}^{S} b_s \sum_{j=0}^{s-1} (\tau^K(\tau-1))^{j+1}\left((\bar{A}^j)(e_1 + \tau^{K+1}e_r)\right)_s.$$

Step (ii): Asymptotic expansion of τ for each of the two stability regions $\mathcal{R}_1^{\mathrm{PRK}}$ and $\mathcal{R}_2^{\mathrm{PRK}}$ Let us first consider the region $\mathcal{R}_1^{\mathrm{PRK}}$, in which case τ is close to one. We propose an asymptotic expansion for τ as a function of $z = \delta t / \Delta t$, and look for those values of z for which (59) is satisfied. We expand $\tau(\theta)$ as follows:

$$\tau(\theta) = 1 + C_1(\theta)z + C_2(\theta)z^2 + h.o.t \qquad 0 \le \theta \le 2\pi$$

By means of an application of the binomial theorem on both τ^{K+1} and $(\tau^K(\tau-1))^{j+1}$, we get

$$\tau^{K+1} = 1 + C_1(K+1)z + (K+1)\left(C_2 + \frac{(K)}{2}C_1^2\right)z^2 + O(z^3),$$

$$(\tau^K(\tau-1))^j = C_1^j z^j + jC_1^{j-1}z^{j-1}\left((KC_1^2 + C_2)z^2 + \left(2KC_1C_2 - \frac{1}{2}KC_1^3 + \frac{1}{2}K^2C_1^3\right)z^3\right) + O(z^4),$$

where we have momentarily suppressed dependence on θ. Then, $F_s(z)$ can be expanded as

$$F_s(z) = \sum_{j=0}^{s-1}(\tau^K(\tau-1))^j\left[\left(\frac{\bar{C}}{z} - (K+1)\right)^j A^j(e_1 + e_r\tau^{K+1})\right]_s,$$

$$= \sum_{j=0}^{s-1}\left(C_1^j + jC_1^{j-1}z(KC_1^2 + C_2)\right)\left[\left(\bar{C}^j + j\bar{C}^{j-1}z(-(K+1))A^j\right)(e_1 + e_r(1 + C_1(K+1)z))\right]_s + O(z^2)$$

where we have introduced $\bar{C}$ as $\mathrm{diag}(0, c_2, \ldots, c_s)$ and applied the binomial theorem to expand $\bar{A}^j$. This allows us to expand σ_S as follows:

$$\sigma_S = 1 + \sum_{s=1}^{S}b_s\sum_{l=1}^{s-1}C_1^j\left[\left(\bar{C}^jA^je_1\right)_s + \left(\bar{C}^jA^je_r\right)_s\right] + O(z).$$

The coefficients C_1, C_2 can be determined by solving the equation $\sigma_S = exp(\imath\theta)$ and matching powers of $(\delta t/\Delta t)$.

In a similar way, this approach can be applied to the region $\mathcal{R}_{2,\mathrm{PRK}}$, where we expand $\tau(\theta)$ as:

$$\tau(\theta) = C_1'(\theta)z^{1/K} + C_2'(\theta)z^{2/K} + h.o.t.$$

We now derive the expressions

$$\tau^{K+1} = C_1'^{K+1}z^{K+1} + (K+1)C_1'^K C_2 z^{K+2} + O(z^{K+3}),$$

$$\left(\tau^K(\tau-1)\right)^j = (-1)^j C_1'^{jK}z^{jK} + (-1)^{j-1}z^{jK+1}\left(C_1'^{jK+1} - jKC_1'^{jK-1}C_2'\right),$$

from which we obtain

$$F_s = \sum_{j=0}^{s-1}(-1)^j C_1'^{jK-1}\left(C_1' - z\left(C_1'^2 - jKC_2'\right)\right)\bar{C}^{j-1}\left(\bar{C} - j(K+1)z\right)(A^je_1)_s + O(z^2)$$

$$= \sum_{j=0}^{s-1}C_1'^{jK}\bar{C}^j(A^je_1)_s + O(z),$$

which implies that σ_S reads:

$$\sigma_S = \sum_{s=1}^{S}b_s\sum_{l=0}^{s-1}C_1'^{lK}\bar{C}^l(A^le_1)_s + O(z).$$

We conclude by giving two concrete examples.

Example 4 (Parametrization of stability regions of PRK2). *Let us first consider the second order projective Runge–Kutta method PRK2. We start with the region $\mathcal{R}_{2,\mathrm{PRK}}$. To this end, we have to solve the following equation for C_1'*

$$\frac{1}{2}(C_1')^{2K} = e^{\imath\theta},$$

which yields the roots:

$$C_1' = \sqrt{2}^{1/K} \exp(\imath(\theta/K + 2j\pi/K)) \quad j = 1,\ldots,K-1.$$

Hence, $\tau(\theta)$ can be written as

$$\tau = \sqrt{2}^{1/K} \exp \imath(\theta + 2j\pi/K) \left(\frac{\delta t}{\Delta t}\right)^{1/K}.$$

*For the stability region $\mathcal{R}_{1,\mathrm{PRK}}$, we use equation (**??**) to determine the coefficient C_1,*

$$1 - e^{\imath\theta} + C_1 + \frac{1}{2}C_1^2 = 0.$$

which gives rise to the solution:

$$C_1 = -1 \pm \sqrt[4]{5 - 4\cos\theta} \arg(-1 + 2\cos\theta + 2\imath\sin\theta).$$

Thus, this part of the stability region is defined by:

$$\tau = 1 - \left(\frac{\delta t}{\Delta t}\right) \pm \sqrt[4]{5 - 4\cos\theta} \arg(-1 + 2\cos\theta + 2\imath\sin\theta)\left(\frac{\delta t}{\Delta t}\right) + O\left(\left(\frac{\delta t}{\Delta t}\right)^2\right)$$

Example 5 (Parametrization of stability regions of PRK4). *The derivation is very similar to the second order case. Now, we will determine the stability region around zero. So we will expand the amplification factor of the inner integrator τ in powers of $(\tau/\Delta t)^{1/K}$:*

$$\tau = C_1' \left(\frac{\delta t}{\Delta t}\right)^{1/K} + C_2' \left(\frac{\delta t}{\Delta t}\right)^{2/K},$$

and substitute the latter into the stability polynomial equation. To find the parametrization of the stability region, we set $\sigma = e^{\imath\theta}$ since we are looking for the values of τ that result in a amplification factor $|\sigma| = 1$. Next, we have to match the powers of $(\delta t/\Delta t)^{1/K}$ on both sides of the equations. This yields the following fourth order polynomial in C_1^K:

$$\frac{1}{24}\left(C_1'^K\right)^4 - \frac{1}{12}\left(C_1'^K\right)^3 + \frac{1}{6}\left(C_1'^K\right)^2 - \frac{1}{6}C_1'^K - e^{\imath\theta} = 0$$

This equation can be solved by using Ferrari's method [47]. First, we have to convert this polynomial into a so called depressed quartic, by performing a change of variables: $C_1'^K = x + 1/2$, which reduces the quartic to :

$$p_1(x) = x^4 + \frac{5}{2}x^2 - x - \frac{19}{16} - 24e^{\imath\theta} = 0$$

The latter can be factored into quadratic polynomials:

$$p_1(x) = (x^2 + px + q)(x^2 + rx + s)$$

which results in the resolvent cubic polynomial in $P = p^2$:

$$P^3 + 5P^2 + (11 + 96e^{\imath\theta})P - 1 = 0$$

This can be solved by performing again a change of variables : $P = t - 5/3$ to reduce the polynomial to a depressed cubic:

$$t^3 + \left(\frac{8}{3} + 96e^{i\theta}\right)t - \frac{272}{27} - 160e^{i\theta} = 0,$$

followed by Viéta's substitution: $t = w - \dfrac{8/3 + 96e^{i\theta}}{3w}$ which finally yields with a quadratic polynomial in w^3:

$$w^6 + \left(-160e^{i\theta} - \frac{272}{27}\right)w^3 - 32768e^{3i\theta} - \frac{512}{729} - \frac{2048}{27}e^{i\theta} - \frac{8192}{3}e^{2i\theta} = 0,$$

which yields:

$$w^3 = 80e^{i\theta} + \frac{136}{27} \pm \frac{8}{9}\sqrt{41472e^{3i\theta} + 11556e^{2i\theta} + 1116e^{i\theta} + 33},$$

and hence, we can calculate a possible root of the sextic equation. This implies that w_1 is a possible root:

$$w_1 = \frac{2}{3}\sqrt[3]{270e^{i\theta} + 17 + 3\sqrt{41472e^{3i\theta} + 11556e^{2i\theta} + 1116e^{i\theta} + 33}}$$

which implies that the following expression for t is a root of equation:

$$t = w_1 - \frac{8/3 + 96e^{i\theta}}{3w_1}.$$

Finally, we can calculate C_1' by substituting the expression for t back into the equation for P and using the fact that $P = t - 5/3$ and $C_1'^K = x + 1/2$.

$$C_1' = \sqrt[\kappa]{\frac{\pm\sqrt{P} \pm \sqrt{P - 2\left(\frac{5}{2} + P \pm \frac{1}{\sqrt{P}}\right)}}{2} + \frac{1}{2}}$$

A similar procedure can be followed to derive an expression for C_1 to determine the region $\mathcal{R}_{1,\mathrm{PRK}}$.