

Hypocoercivity and diffusion limit of a finite volume scheme for linear kinetic equations

Marianne Bessemoulin-Chatard, Maxime Herda, Thomas Rey

▶ **To cite this version:**

Marianne Bessemoulin-Chatard, Maxime Herda, Thomas Rey. Hypocoercivity and diffusion limit of a finite volume scheme for linear kinetic equations. Mathematics of Computation, 2020, 89, pp.1093-1133. 10.1090/mcom/3490 . hal-01957832v2

HAL Id: hal-01957832
https://hal.science/hal-01957832v2

Submitted on 16 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HYPOCOERCIVITY AND DIFFUSION LIMIT OF A FINITE VOLUME SCHEME FOR LINEAR KINETIC EQUATIONS

MARIANNE BESSEMOULIN-CHATARD, MAXIME HERDA, AND THOMAS REY

Abstract. In this article, we are interested in the asymptotic analysis of a finite volume scheme for one dimensional linear kinetic equations, with either Fokker-Planck or linearized BGK collision operator. Thanks to appropriate uniform estimates, we establish that the proposed scheme is Asymptotic-Preserving in the diffusive limit. Moreover, we adapt to the discrete framework the hypocoercivity method proposed by [J. Dolbeault, C. Mouhot and C. Schmeiser, *Trans. Amer. Math. Soc.*, 367, 6 (2015)] to prove the exponential return to equilibrium of the approximate solution. We obtain decay rates that are bounded uniformly in the diffusive limit. Finally, we present an efficient implementation of the proposed numerical schemes, and perform numerous numerical simulations assessing their accuracy and efficiency in capturing the correct asymptotic behaviors of the models.

Keywords: Kinetic equations, finite volume methods, hypocoercivity, diffusion limit, asymptotic-preserving schemes.

2010 Mathematics Subject Classification: 82B40, 65M08, 65M12.

Contents

1. Introduction 2
2. The continuous setting 6
2.1. Uniform estimates 7
2.2. Diffusive limit 8
2.3. Hypocoercivity 8
3. The fully discrete setting 10
3.1. Notations 10
3.2. Presentation of the schemes 13
3.3. Uniform estimates 14
3.4. Asymptotic-preserving property 16
3.5. Numerical hypocoercivity 18
4. Implementation 22
4.1. Micro-macro formulations of the schemes 23
4.2. Explicit matrix formulation 25
5. Numerical simulations 26
5.1. Numerical investigation of the AP Property 26
5.2. The Fokker-Planck case 27
5.3. The BGK case 29
5.4. The Non-Gaussian BGK case 33

MBC was partially funded by the Centre Henri Lebesgue (ANR-11-LABX-0020-01) and ANR Project MoHyCon (ANR-17-CE40-0027-01).

MH was partially funded by ANR Project MoHyCon (ANR-17-CE40-0027-01).

TR were partially funded by Labex CEMPI (ANR-11-LABX-0007-01) and ANR Project MoHyCon (ANR-17-CE40-0027-01).

6. Appendix. On discrete Poincaré inequalities 33
References 36

1. Introduction

The linear kinetic equation. Many engineering or biological problems involve fluid-like systems in transitional regimes: micro-electro-mechanical systems, space shuttle reentry, powder and grains in storage silos, bacteria colonizing a medium, *etc.* As a consequence of nonequilibrium behaviors, the macroscopic descriptions can break down, and a kinetic model may be needed to depict accurately the system. In this contribution, we are interested in the numerical approximation of a prototypical model for such systems. It describes the evolution of a particle distribution function $f^\varepsilon = f^\varepsilon(t,x,v)$, for time $t \geq 0$, at position $x \in \mathbb{T}$ the one dimensional torus, and velocity $v \in \mathbb{R}$, solution to the initial value problem

$$\begin{cases} \varepsilon \dfrac{\partial f^\varepsilon}{\partial t} + v \dfrac{\partial f^\varepsilon}{\partial x} \;=\; \dfrac{1}{\varepsilon}\, \mathcal{Q}(f^\varepsilon), \\ \\ f^\varepsilon(0,x,v) = f_0(x,v) \geq 0\,. \end{cases} \tag{1.1}$$

The *collision* operator $\mathcal{Q}$ describes a microscopic "collision process" acting only on the velocity variable v and preserving the zeroth order moment of the distribution f^ε, namely the mass or total number of particles. The small scaling parameter $\varepsilon > 0$ is the ratio between the mean free path of particles and the length scale of observation. In the context of rarefied gas dynamics it corresponds to the dimensionless *Knudsen* number. The presence of the factor ε in front of the time derivative accounts for the fact that the system is observed on long time scales. The system is said to be in the kinetic regime if $\varepsilon \sim 1$ and in the diffusive regime if $\varepsilon \ll 1$.

Throughout this paper, the collision operator will be either of linear Fokker-Planck type:

$$\mathcal{Q}_{FP}(f)(v) = \partial_v \left(\partial_v f + v f\right) \tag{1.2}$$

or of linearized BGK [7] / relaxation type:

$$\mathcal{Q}_{BGK}(f)(v) = \rho M(v) - m_0 f(v), \qquad \rho = \int_{\mathbb{R}} f(v)\,\mathrm{d}v \tag{1.3}$$

where M is a given nonnegative function of v, the so-called *Maxwellian* whose moments are denoted by

$$m_k \;=\; \int_{\mathbb{R}} |v|^k M(v)\,\mathrm{d}v.$$

In particular, m_0 is the mass of M.

Remark 1. *The linear BGK model appears quite naturally as a linearization of the complete, quadratic, Boltzmann equation* [11] *near its global Maxwellian equilibrium M. Indeed, one can write*

$$\begin{aligned} \mathcal{Q}_{BGK}(f)(v) &= M(v) \int_{\mathbb{R}} f(v_*)\,\mathrm{d}v_* - f(v) \int_{\mathbb{R}} M(v_*)\,\mathrm{d}v_* \\ &= \int_{\mathbb{R}} \left[f_*\, M - f\, M_*\right] \mathrm{d}v_*, \end{aligned}$$

where the usual shorthand notations $\psi := \psi(v)$ and $\psi_* := \psi(v_*)$ were used.

In the case of the BGK operator, every result in this article applies to Maxwellians M that are even functions of v and which have finite moments $m_k < +\infty$ up to $k = 4$ (see Section 5.4 for an illustration). However, in order to unify the analysis of both collision operators, we will focus in the rest of the presentation on the classical Boltzmann's Maxwellian, namely the centered reduced Gaussian function

$$M(v) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{|v|^2}{2}\right).$$

In particular, $m_0 = 1$ and $\mathcal{Q}_{BGK}(f) = \rho M - f$.

Remark 2. *Observe that such a Maxwellian distribution verifies the elementary identity*

$$M'(v) = -vM(v). \tag{1.4}$$

Hence, one can rewrite the Fokker-Planck operator in the following gradient form

$$\mathcal{Q}_{FP}(f) = \partial_v \left(M\, \partial_v \left(\frac{f}{M} \right) \right). \tag{1.5}$$

The motivations in this paper are twofold and concern the design and analysis of a numerical scheme which is accurate in two asymptotics of (1.1). The first asymptotics is the diffusion limit $\varepsilon \to 0$ and the second is the long-time asymptotics $t \to \infty$.

Diffusion limit. A natural mathematical problem concerning kinetic equations such as (1.1) is the study of the diffusion / parabolic limit $\varepsilon \to 0$. Thanks to the choice of time scale, it is possible to capture nontrivial macroscopic dynamics at the limit. In order to identify the limit, one can multiply the equation (1.1) by $(1, v)$ and integrating in the velocity space to get

$$\begin{cases} \partial_t \rho^\varepsilon + \partial_x j^\varepsilon = 0, \\ \varepsilon^2 \partial_t j^\varepsilon + m_2 \partial_x \rho^\varepsilon + \partial_x S^\varepsilon = -j^\varepsilon, \end{cases} \tag{1.6}$$

where the moments are defined by

$$\rho^\varepsilon := \int_{\mathbb{R}} f^\varepsilon \, \mathrm{d}v, \quad j^\varepsilon := \frac{1}{\varepsilon} \int_{\mathbb{R}} f^\varepsilon \, v \, \mathrm{d}v, \quad S^\varepsilon := \int_{\mathbb{R}} (v^2 - m_2) f^\varepsilon \, \mathrm{d}v.$$

Then, by formally taking limits $\varepsilon \to 0$ in (1.1) and (1.6), one obtains $f^\varepsilon \to f = \rho M$ which implies $S^\varepsilon \to 0$ and the second equation of (1.6) yields $j^\varepsilon \to -m_2 \partial_x \rho$. Hence, the limit density ρ is solution to the linear heat equation

$$\frac{\partial \rho}{\partial t} - \partial_x (m_2 \partial_x \rho) = 0. \tag{1.7}$$

In the literature, the first results concerning the approximation of kinetic models by diffusion equations have been proposed in [4, 41]. Justifications of asymptotic expansion of the solutions in power of ε can be found in [3] for the neutron transport, in [18] for the Fokker-Planck equation, in [2] for the radiative transfer equation, and in [49] for linear semiconductor Boltzmann equation. In [43], the authors studied the diffusive limit of generalized two-velocity kinetic models. A large class of linear collision operators was considered in [17], and approximation of kinetic equations by diffusion is justified by homogenization in [29]. We also refer to [8, Chapter 2] for an overview of classical hydrodynamic and parabolic limits of kinetic equations. Extensions to more general models with fractional diffusion limits were more recently obtained in [45].

There exists a dense literature describing numerical schemes which enjoy the property of being stable in the diffusion limit, and converge to a numerical discretization of the macroscopic model. These schemes fall down in the so-called Asymptotic-Preserving (AP) framework and are compatible with the kinetic and asymptotic regime of the equation. The principle of AP schemes can be roughly summarized as the commutative diagram presented in Fig. 1.

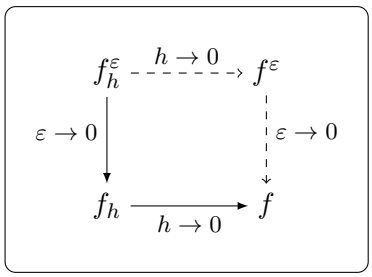


FIGURE 1. The AP diagram (h denotes the size of the discretization)

In the framework of kinetic equations, AP schemes first appeared 20 years ago, on the one hand for kinetic semiconductors models in two papers due to Klar [39] for the linear BGK equation and to Schmeiser and Zwirchmayr [52] for the linear Boltzmann equation, and on the other hand for relaxation models (*e.g.* two velocities Goldstein-Taylor) in the article of Jin, Pareschi, and Toscani [36]. Two recent review papers [35, 19] contain an almost up to date bibliography on numerical methods for collisional kinetic equations of type (1.1) and AP schemes. For even more recent works on AP schemes, one can consult these two papers [20, 15] on particle methods and the references therein.

Nevertheless, most of these works contain only formal proofs of the AP properties, the very important solid lines of Fig. 1. Here are some instances of AP schemes with proofs: the micro-macro approach of Lemou and Mieussens [42], for either discrete velocity models (Goldstein-Taylor-like) or linear scattering model (continuous in velocity), contain some partial proofs. These proofs were later completed by Liu and Mieussens in [44] with a very detailed stability analysis. Note that the numerical scheme for dealing with the kinetic equation is semi-implicit in time, but the limiting $\varepsilon \to 0$ scheme is explicit, yielding a very restrictive parabolic CFL condition on the time step. For the anomalous diffusion limit, let us also mention the work of Crouseilles, Hivert, and Lemou [16] where an AP scheme is built, with a proof of convergence (semi-discrete in time), without rate of convergence towards the equilibrium.

Hypocoercivity and large time behavior. A second classical problem in the asymptotic analysis of kinetic equations is the large time behavior of solutions. On the class of models under consideration, there are several ways to show that there are constants $\kappa > 0$ and $C \geq 1$ such that the solution satisfies

$$\|f^\varepsilon(t) - \mu_f M\|_{\mathcal{X}} \leq C\|f_0 - \mu_f M\|_{\mathcal{X}} e^{-\kappa t}, \tag{1.8}$$

where

$$\mu_f = \iint_{\mathbb{T}\times\mathbb{R}} f_0 \,\mathrm{d}x\,\mathrm{d}v\,,$$

and $\mathcal{X}$ is some appropriate functional space. In the diffusive scaling under consideration, the constants C and κ can be chosen independently of ε as ε tends to zero. In the case of the kinetic Fokker-Planck operator, the model is simple enough so that it is actually possible to compute explicitly the fundamental solution and derive such an estimate.

However, in recent years, more systematic and robust methods have been developed in order to show (1.8). They are called *hypocoercivity* methods and can usually be seen as entropy methods [37, 51], in the sense that the goal is to design a functional $H(f^\varepsilon)$, called the (modified) entropy, which is dissipated along the solutions. From there a control of the entropy by its dissipation provides decay of the former with explicit rates. Hypocoercivity methods deal with the fact that

for kinetic equations such as (1.1), dissipation / relaxation apparently occurs only in the velocity variable. Nonetheless, thanks to the mixing properties of the transport operator $v\partial_x$ in the phase space $x - v$, one can recover dissipation along x and v. The denomination hypocoercivity is to be related with the fact that if (1.8) held with $C = 1$ then the operator $v\partial_x - \mathcal{Q}(\cdot)$ would be coercive in the usual sense.

A comprehensive introduction on hypocoercivity methods may be found in the lectures of Hérau [31]. The first hypocoercivity method introduced by Hérau and Nier for kinetic Fokker-Planck equations [32] is very similar to the method of commutators for proving hypoellipticity going back to the seminal work of Hörmander [33]. This method has been successfully applied to non-hypoelliptic equations (without regularization effects), such as linear Boltzmann or BGK equations, by Hérau [30], and Mouhot and Neumann [46]. A general abstract framework for this method of commutators has then been given by Villani in [53]. The drawback of these hypocoercivity methods is that they often require regularity on the initial datum and involve estimates in equivalent weighted H^1 norms leading to sometimes tedious computations. In [30], the technique is adapted to start from only weighted L^2 initial datum, but the structure of the functional is similar.

In a recent work [21], Dolbeault, Mouhot and Schmeiser established a new hypocoercivity method in a general abstract framework. It is based on the definition of a modified entropy taking the form of an equivalent weighted L^2 norm involving macroscopic quantities (*i.e.* integrated in velocity). In this work we focus on this particular hypocoercivity method, and we show that it is well-suited for an adaptation in the discrete framework.

Concerning the numerical methods preserving large-time behaviors of solutions, several efficient numerical schemes have been developed for homogeneous kinetic equations. A full discrete finite difference scheme for the homogeneous Fokker-Planck equation was built in the pioneering work of Chang and Cooper [14]. More recently, several schemes preserving the exponential trend to equilibrium have been proposed for nonlinear degenerate parabolic equations (see for example [5, 10, 12, 28]). In [47], the question of describing correctly the equilibrium state of several nonlinear diffusion and kinetic equations is addressed together with that of the order of the schemes. Let us also mention [24, 13] where the case of non-homogeneous Dirichlet boundary conditions are dealt with.

In the case of inhomogeneous kinetic equations, only few results are available, as the understanding of hypocoercive structures is quite recent. The case of the very simple Kolmogorov equation has been investigated in [48, 26, 27]. In [26], a time-splitting technique based on self-similarity properties is used for solutions that decay polynomially in time. In [48, 27], the H^1 method *à la* Hérau and Villani has been adapted for respectively a finite difference and a finite element schemes. Lastly, let us finish with the recent paper of Dujardin, Hérau and Lafitte [22] where the same H^1 method is adapted to a finite difference scheme for the kinetic Fokker-Planck equation (1.1)–(1.2).

Main results. In this contribution we design implicit in time finite volume schemes for (1.1) with collision operators (1.2) and (1.3). We perform their numerical asymptotic analysis and are able to show the following properties.

- Our schemes are AP (Theorem 1) in the diffusion limit $\varepsilon \to 0$. Solutions converge to solutions of a finite difference scheme solving the heat equation (1.7). Provided that the scheme is implemented correctly (see Section 4) one can take $\varepsilon = 0$ in the scheme. In particular the linear system that has to be solved at each time iteration has coefficients and condition number that are uniformly bounded with respect to ε.
- The solutions of our schemes preserve the decay of a discrete analogue of the modified entropy of Dolbeault, Mouhot and Schmeiser [21]. Consequently, we are able to prove

the discrete equivalent of the exponential decay estimate (1.8) (Theorem 2). The decay rates are bounded uniformly in the diffusion limit.

Some comments are in order. On the first hand, let us mention that an essential identity which allows to prove both properties in the Fokker-Planck case is a discrete equivalent of (1.4). By choosing the framework of finite volume schemes we can introduce a discrete Maxwellian at cell centers and interfaces of the mesh in velocity. Enforcing the discrete equivalent of (1.4) just relates these two quantities and leaves a large choice in the actual definition of the discrete Maxwellian.

On the second hand, by choosing to adapt the L^2 hypocoercivity method rather than the H^1 method, we avoid the burden of estimating derivatives of the solution and dealing with commutators coming from the choice of discrete derivative operators. Thus, the numerical analysis turns out to be almost identical to what is done in the continuous case. For this reason, we are confident that this method could successfully be applied for the asymptotic analysis of other numerical schemes for hypocoercive equations.

Plan of the paper. The outline of this article is as follows. In Section 2, we briefly present the results in the continuous setting. We start by establishing some uniform estimates, which are the key point to study both the diffusive limit and the exponential return to equilibrium. Then we state the convergence of the kinetic equation towards the heat equation as ε tends to 0. To conclude this section, we prove the L^2-hypocoercivity of both Fokker-Planck and linearized BGK operators, by introducing a modified entropy functional which is a slight simplification of the original version proposed by Dolbeault, Mouhot and Schmeiser [21]. In Section 3, we adapt these results to the discrete framework. The considered numerical schemes are implicit in time, and finite volume in space and velocity. We discretize carefully the Fokker-Planck fluxes, in order to obtain a discrete version of the L^2 entropy estimate, which is the starting point to establish the needed discrete uniform estimates. Then our two main results are stated in Theorem 1 (asymptotic-preserving property in the diffusive limit) and in Theorem 2 (numerical hypocoercivity). Section 4 is then devoted to the practical implementation of our method. We present a discretization equivalent to the proposed schemes, based on a perturbative micro-macro formulation, and we write the obtained linear systems in explicit matrix forms for sake of completeness. Finally in Section 5, we present several numerical experiments which demonstrate the efficiency of our schemes. After investigating numerically the asymptotic-preserving property, we finally study the trend to equilibrium for both Fokker-Planck and linearized BGK operators.

Acknowledgments. The authors would like to thank G. Dujardin, F. Hérau and P. Lafitte for fruitful discussions during the conception of this article.

2. The continuous setting

For convenience, let us introduce the measures $\mathrm{d}M := M(v)\mathrm{d}v$,

$$\mathrm{d}\gamma = \gamma(v)\mathrm{d}v := \tfrac{\mathrm{d}v}{M(v)}$$

and $L^2(\mathrm{d}M)$ (*resp.* $L^2(\mathrm{d}\gamma)$) the spaces of square integrable functions against the measure $\mathrm{d}M$ (*resp.* $\mathrm{d}\gamma$). A straightforward consequence of the Gaussian Poincaré inequality is that for any $f \in L^2(\mathrm{d}\gamma)$, one has

$$\|f - \rho M\|_{L^2(\mathrm{d}\gamma)} \leq \|\partial_v \left(\tfrac{f}{M}\right)\|_{L^2(\mathrm{d}M)}, \tag{2.1}$$

where $\rho = \int f \mathrm{d}v$.

2.1. **Uniform estimates.** With an initial data in $L^2(\mathrm{d}x\mathrm{d}\gamma)$, there is a unique solution to (1.1), with either Fokker-Planck or BGK collision operators (see *e.g.* [1]). This solution conserves mass and nonnegativity. Additionally one has the following basic "entropy" estimate.

Lemma 1. *Let f^ε be the solution of* (1.1) *with initial data $f_0 \in L^2(\mathrm{d}x\mathrm{d}\gamma)$. Then for every $t \geq 0$,*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|f^\varepsilon(t)\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)} + \frac{1}{\varepsilon^2}\|f^\varepsilon(t)-\rho^\varepsilon(t)M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)} \leq 0\,. \tag{2.2}$$

In particular one has

$$\max\left(\|f^\varepsilon\|^2_{L^\infty_t L^2(\mathrm{d}x\mathrm{d}\gamma)}, \frac{2}{\varepsilon^2}\|f^\varepsilon-\rho^\varepsilon M\|^2_{L^2_t L^2(\mathrm{d}x\mathrm{d}\gamma)}\right) \leq \|f_0\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,. \tag{2.3}$$

Proof. Estimate (2.2) is obtained by multiplying (1.1) with f^ε/M and integrating in x and v yielding

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|f^\varepsilon\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)} + \frac{1}{2\varepsilon}\iint \partial_x\left(v\left(\frac{f^\varepsilon}{M}\right)^2\right)\mathrm{d}x\mathrm{d}v = \frac{1}{\varepsilon^2}\iint \mathcal{Q}(f^\varepsilon)\frac{f^\varepsilon}{M}\mathrm{d}x\,\mathrm{d}v\,.$$

The second term of the left hand side cancels.
In the case of the BGK operator, since $\int \mathcal{Q}_{BGK}(f^\varepsilon)\mathrm{d}v = 0$, one has

$$\iint \mathcal{Q}_{BGK}(f^\varepsilon)\frac{f^\varepsilon}{M}\mathrm{d}x\,\mathrm{d}v = \iint \mathcal{Q}_{BGK}(f^\varepsilon)\left(\frac{f^\varepsilon}{M}-\rho^\varepsilon\right)\mathrm{d}x\,\mathrm{d}v = -\|f^\varepsilon-\rho^\varepsilon M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,.$$

In the case of the Fokker-Planck operator, an integration by parts in v and the Poincaré inequality (2.1) yields

$$\iint \mathcal{Q}_{FP}(f^\varepsilon)\frac{f^\varepsilon}{M}\mathrm{d}x\,\mathrm{d}v = -\|\partial_v\left(\frac{f^\varepsilon}{M}\right)\|^2_{L^2(\mathrm{d}x\mathrm{d}M)} \leq -\|f^\varepsilon-\rho^\varepsilon M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}.$$

□

From the entropy estimate, one can gather many estimates on the moments.

Lemma 2 (Moments estimates)**.** *Under the hypotheses of Lemma 1, the moments ρ^ε, j^ε and S^ε satisfy equations* (1.6) *and the following estimates:*

$$\|\rho^\varepsilon(t)\|_{L^2_x} \leq \|f^\varepsilon(t)\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,, \tag{2.4}$$

$$\varepsilon\,\|j^\varepsilon(t)\|_{L^2_x} \leq m_2^{1/2}\,\|f^\varepsilon(t)\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)} \tag{2.5}$$

$$\varepsilon\,\|j^\varepsilon(t)\|_{L^2_x} \leq m_2^{1/2}\,\|f^\varepsilon(t)-\rho^\varepsilon(t)\,M\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,, \tag{2.6}$$

$$\|S^\varepsilon(t)\|_{L^2_x} \leq (m_4-m_2^2)^{1/2}\,\|f^\varepsilon(t)-\rho^\varepsilon(t)\,M\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,. \tag{2.7}$$

In particular, there is a constant $C>0$ depending only on m_2 and m_4 such that

$$\max\left(\|\rho^\varepsilon\|_{L^\infty_t L^2_x},\, \varepsilon\,\|j^\varepsilon\|_{L^\infty_t L^2_x},\, \|j^\varepsilon\|_{L^2_t L^2_x},\, \frac{1}{\varepsilon}\,\|S^\varepsilon\|_{L^2_t L^2_x}\right) \leq C\,\|f_0\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,. \tag{2.8}$$

Proof. In order to estimate the moments, let us observe that one can rewrite the macroscopic quantities as

$$\begin{aligned}\rho^\varepsilon &= \int M^{1/2}(f^\varepsilon M^{-1/2})\,\mathrm{d}v,\\ \varepsilon\, j^\varepsilon &= \int (vM^{1/2})(f^\varepsilon M^{-1/2})\,\mathrm{d}v \;=\; \int (vM^{1/2})(f^\varepsilon - \rho^\varepsilon M)M^{-1/2}\,\mathrm{d}v,\\ S^\varepsilon &= \int (v^2-m_2)M^{1/2}(f^\varepsilon-\rho^\varepsilon M)M^{-1/2}\,\mathrm{d}v,\end{aligned}$$

by symmetry of M for j^ε, and by definition of m_2 for S^ε. Then one can use Cauchy-Schwarz inequalities to get the desired estimates. Estimate (2.8) is then a consequence of (2.3). Finally, integrate (1.1) in v against 1 and v to get (1.6). □

2.2. **Diffusive limit.**

Proposition 1. *Let $(f^\varepsilon)_{\varepsilon\geq 0}$ be a sequence of solutions of (1.1) (with either Fokker-Planck or BGK collision operator) with initial data $f_0 \in L^2(\mathrm{d}x\mathrm{d}\gamma)$. Then when $\varepsilon \to 0$, one has*

$$f^\varepsilon(t,x,v) \to f(t,x,v) \;=\; \rho(t,x)\,M(v) \text{ weakly in } L^2_{t,loc}L^2(\mathrm{d}x\mathrm{d}\gamma)$$

and ρ solves the heat equation

$$\partial_t\rho - m_2\,\partial^2_{xx}\rho \;=\; 0\,,\quad \rho(0,x) \;=\; \int f_0(x,v)\,\mathrm{d}v.$$

Proof. Using that $f^\varepsilon - \rho^\varepsilon M$ is uniformly $O(\varepsilon)$ in $L^2(0,\infty;L^2(\mathrm{d}x\mathrm{d}\gamma))$ and that ρ^ε is uniformly bounded in $L^\infty_t L^2_x$, we get the existence of the limit, up to extraction of a subsequence. Then, by combining both equations of (1.6) one gets

$$\partial_t\rho^\varepsilon - m_2\partial^2_{xx}\rho^\varepsilon \;=\; \varepsilon^2\partial_t j^\varepsilon + \partial_x S^\varepsilon$$

in the sense of distributions. Taking limits thanks to the estimates (2.8) one recovers the heat equation. By uniqueness of the limit the whole sequence converges. □

2.3. **Hypocoercivity.** Let $\mathcal{H}$ be the subspace of the Hilbert space $L^2(\mathrm{d}x\mathrm{d}\gamma)$ composed of the zero mean functions with canonical scalar product. The space $\mathcal{H}$ coincides with the orthogonal space of the nullspace of the unbounded operator $L_\varepsilon = \frac{1}{\varepsilon}v\partial_x - \frac{1}{\varepsilon^2}\mathcal{Q}(\cdot)$ in both the Fokker-Planck and BGK cases. One can show thanks to the method developed in [21] that the latter operator is hypocoercive (see [53]) in $\mathcal{H}$ uniformly for small ε. More precisely, one has the following result.

Proposition 2. *There are constants $C \geq 1$ and $\kappa > 0$ such that for all $\varepsilon \in (0,1)$ and all initial data $f_0 \in L^2(\mathrm{d}x\mathrm{d}\gamma)$, the solution f^ε of (1.1) satisfies*

$$\|f^\varepsilon(t) - \mu_f M\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)} \;\leq\; C\,\|f_0 - \mu_f M\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,e^{-\kappa\,t}.$$

The rest of the section is devoted to the proof of Proposition 2. We need some intermediate results stated in the following lemmas. Up to changing f_0 by $f_0 - \mu_f M$ and correspondingly f^ε and ρ^ε by respectively $f^\varepsilon - \mu_f M$ and $\rho^\varepsilon - \mu_f$, we can assume from now on in this section that

$$\mu_f \;=\; \iint f_0 \mathrm{d}x\mathrm{d}v \;=\; 0\,.$$

It follows that f^ε and ρ^ε are mean-free.

In order to prove Proposition 2, we introduce the "modified entropy functional" *à la* Dolbeault, Mouhot and Schmeiser. It is a slight simplification of the original version of [21] using the fact that we work on a bounded space domain. It reads

$$H(f^\varepsilon(t)) = \frac{1}{2}\|f^\varepsilon(t)\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)} + \eta\,\varepsilon^2\,\langle j^\varepsilon, \partial_x\phi^\varepsilon\rangle_{L^2_x}\,, \tag{2.9}$$

where $\phi^\varepsilon(t,x)$ is the solution of the Poisson equation

$$-\partial^2_{xx}\phi^\varepsilon \;=\; \rho^\varepsilon\,, \qquad \int_{\mathbb{T}}\phi^\varepsilon \mathrm{d}x \;=\; 0$$

and $\eta > 0$ is a small parameter to be chosen.

Lemma 3. *The function ϕ^ε satisfies for all $t \geq 0$*

$$\begin{aligned} \|\partial_x\phi^\varepsilon(t)\|_{L^2_x} &\leq C_P\,\|\rho^\varepsilon(t)\|_{L^2_x}, \\ \|\partial_t\partial_x\phi^\varepsilon(t)\|_{L^2_x} &\leq \|j^\varepsilon(t)\|_{L^2_x}, \end{aligned}$$

where $C_P = 1/2\pi$ is the Poincaré constant of $\mathbb{T}$.

Proof. The first estimate is obtained by testing the Poisson equation against ϕ^ε and applying the Poincaré inequality

$$\|\partial_x\phi^\varepsilon\|^2_{L^2_x} \;=\; \left\langle -\partial^2_{xx}\phi^\varepsilon, \phi^\varepsilon\right\rangle_{L^2_x} \;\leq\; \|\rho^\varepsilon\|_{L^2_x}\,\|\phi^\varepsilon\|_{L^2_x} \;\leq\; C_P\|\rho^\varepsilon\|_{L^2_x}\,\|\partial_x\phi^\varepsilon\|_{L^2_x}\,.$$

For the second estimate, first take the time derivative of the Poisson equation and use the continuity equation $\partial_t\rho^\varepsilon + \partial_x j^\varepsilon = 0$ to get $-\partial_t\partial^2_{xx}\phi^\varepsilon \;=\; -\partial_x j^\varepsilon$. Then multiply by $\partial_t\phi^\varepsilon$ and integrate to get

$$\|\partial_t\partial_x\phi^\varepsilon\|^2_{L^2_x} \;=\; \langle -\partial_x j^\varepsilon, \partial_t\phi^\varepsilon\rangle = \langle j^\varepsilon, \partial_t\partial_x\phi^\varepsilon\rangle_{L^2_x} \;\leq\; \|j^\varepsilon\|_{L^2_x}\,\|\partial_t\partial_x\phi^\varepsilon\|_{L^2_x}.$$

□

For small enough $\eta > 0$, the square root of the modified entropy actually defines an equivalent norm on $\mathcal{H}$.

Lemma 4 (Equivalent norm)**.** *There is $\eta_1 > 0$ such that for all $\eta \in (0,\eta_1)$, there are positive constants $0 < c_\eta < C_\eta$ such that for all $\varepsilon \in (0,1)$ and $f^\varepsilon \in \mathcal{H}$ one has*

$$\frac{c_\eta}{2}\,\|f^\varepsilon\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)} \;\leq\; H(f^\varepsilon) \;\leq\; \frac{C_\eta}{2}\,\|f^\varepsilon\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,.$$

Proof. The result follows directly from the expression of the entropy and the estimate

$$|\,\langle j^\varepsilon, \partial_x\phi^\varepsilon\rangle_{L^2_x}\,| \;\leq\; \|j^\varepsilon\|_{L^2_x}\,\|\partial_x\phi^\varepsilon\|_{L^2_x} \;\leq\; \frac{1}{\varepsilon}\,C_P\,m_2^{1/2}\,\|f^\varepsilon\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,.$$

□

With the previous lemmas, we can prove the main result of this section.

Proof of Proposition 2. Using the moment equation $\varepsilon^2\partial_t j^\varepsilon + m_2\,\partial_x\rho^\varepsilon + \partial_x S^\varepsilon + j^\varepsilon = 0$, the time derivative of the entropy splits into five terms

$$\frac{\mathrm{d}}{\mathrm{d}t}H(f^\varepsilon(t)) \;=\; T_1^\varepsilon(t) + \eta\,m_2\,T_2^\varepsilon(t) + \eta\,T_3^\varepsilon(t) + \eta\,T_4^\varepsilon(t) + \eta\,T_5^\varepsilon(t)\,.$$

Let us precise and estimate each term:

$$\begin{aligned}
T_1^\varepsilon &:= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|f^\varepsilon\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)} &&\leq -\frac{1}{\varepsilon^2}\|f^\varepsilon-\rho^\varepsilon M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,,\\
T_2^\varepsilon &:= -\langle\partial_x\rho^\varepsilon,\partial_x\phi^\varepsilon\rangle_{L^2_x} = \langle\rho^\varepsilon,\partial^2_{xx}\phi^\varepsilon\rangle_{L^2_x} &&= -\|\rho^\varepsilon\|^2_{L^2_x}\,,\\
T_3^\varepsilon &:= -\langle\partial_x S^\varepsilon,\partial_x\phi^\varepsilon\rangle = -\langle S^\varepsilon,\rho^\varepsilon\rangle &&\leq (m_4-m_2^2)^{1/2}\,\|f^\varepsilon-\rho^\varepsilon M\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,\|\rho^\varepsilon\|_{L^2_x}\,,\\
T_4^\varepsilon &:= -\langle j^\varepsilon,\partial_x\phi^\varepsilon\rangle &&\leq \frac{C_P\,m_2^{1/2}}{\varepsilon}\,\|f^\varepsilon-\rho^\varepsilon M\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,\|\rho^\varepsilon\|_{L^2_x}\,,\\
T_5^\varepsilon &:= \varepsilon^2\,\langle j^\varepsilon,\partial_t\partial_x\phi^\varepsilon\rangle &&\leq m_2\|f^\varepsilon-\rho^\varepsilon M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,,
\end{aligned}$$

where we used the estimates of Lemma 2 and Lemma 3. Combining everything, one has

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}H(f^\varepsilon(t)) + (\frac{1}{\varepsilon^2}-\eta\,m_2)\,\|f^\varepsilon-\rho^\varepsilon M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)} + \eta\,m_2\|\rho^\varepsilon\|^2_{L^2_x}\\
&\qquad\leq \eta\left((m_4-m_2^2)^{1/2}+\frac{C_P\,m_2^{1/2}}{\varepsilon}\right)\|f^\varepsilon-\rho^\varepsilon M\|_{L^2(\mathrm{d}x\mathrm{d}\gamma)}\,\|\rho^\varepsilon\|_{L^2_x}\,.
\end{aligned}$$

It follows that for any $\eta\in(0,\eta_2)$ with $\eta_2 = m_2/(((m_4-m_2^2)^{1/2}+C_P\,m_2^{1/2})^2+m_2^2)$ and $\varepsilon\in(0,1)$, one has

$$\frac{\mathrm{d}}{\mathrm{d}t}H(f^\varepsilon(t)) + K_\eta\,(\|f^\varepsilon-\rho^\varepsilon M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}+\|\rho^\varepsilon M\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}) \leq 0\,,$$

with $K_\eta = \min(1-\eta m_2,\eta m_2)/2$. Then use the triangle inequality, $(a+b)^2\leq 2(a^2+b^2)$ and Lemma 4 to get that

$$\frac{\mathrm{d}}{\mathrm{d}t}H(f^\varepsilon(t)) + \frac{K_\eta}{C_\eta}\,H(f^\varepsilon(t)) \leq 0\,,$$

for all $\eta\in(0,\min(\eta_1,\eta_2))$. Choose some admissible η and set $\kappa = K_\eta/(2C_\eta)$ to conclude. □

3. The fully discrete setting

In this section, our aim is now to adapt the previous results to the discrete setting. To proceed, we first introduce an implicit in time and finite volume in space discretization of (1.1) both for Fokker-Planck and BGK cases. We then prove the discrete counterparts of the results presented in the previous section, namely we study the diffusive limit and the exponential return to equilibrium.

3.1. Notations.

Mesh. We first restrict the velocity domain to a bounded symmetric segment $[-v_\star,v_\star]$, since it is not possible in practice to implement a numerical scheme on an unbounded domain. We consider a primal mesh of this interval composed of $2L$ control volumes arranged symmetrically around $v=0$. We thus get $2L+1$ distinct interface points denoted by $v_{j+\frac12}$ (for consistency with usual finite volume notations) with $j=-L,\dots,L$. In this way

$$v_{-L+1/2} = -v_\star\,,\quad v_{1/2}=0\,,\quad v_{j+1/2}=-v_{-j+1/2}\quad\forall j=0,\dots,L\,.$$

The cells of the primal mesh are given by

$$\mathcal{V}_j := (v_{j-\frac12},v_{j+\frac12})\,,\quad j\in\mathcal{J} := \{-L+1,\dots,L\}.$$

Each cell $\mathcal{V}_j$ has length $\Delta v_j = v_{j+\frac12}-v_{j-\frac12}$ and midpoint v_j. At the level of cells, the symmetry of the velocities reads $v_j = -v_{-j+1}$ for all $j=1,\dots,L$. We also define the dual mesh with cells

$$\mathcal{V}^*_{j+1/2} := (v_j,v_{j+1})\,,\quad j\in\mathcal{J}^* := \{-L,\dots,L\}\,,$$

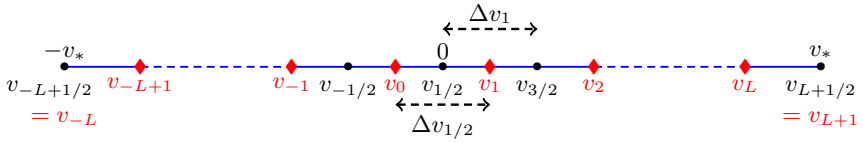


FIGURE 2. Discretization of the velocity domain.

with $v_{-L} := v_{-L+1/2} = -v_\star$ and $v_{L+1} := v_{L+1/2} = v_\star$. The length of the dual cell $\mathcal{V}^*_{j+1/2}$ is $\Delta v_{j+1/2} = v_{j+1} - v_j$. Notations introduced are illustrated in Fig. 2.

In space, we consider a discretization of the torus $\mathbb{T}$ into N subintervals

$$\mathcal{X}_i \ := \ (x_{i-\frac12}, x_{i+\frac12}) \,, \quad i \in \mathcal{I} := \mathbb{Z}/N\mathbb{Z}$$

of length Δx_i and centers x_i. In what follows, we assume that N is odd. Indeed, this assumption is natural to obtain, among others, a discrete Poincaré inequality on the torus with our choice of discrete gradients, as explained later.

The control volumes in phase space are defined by

$$K_{ij} \ := \ \mathcal{X}_i \times \mathcal{V}_j \,, \qquad \forall (i,j) \in \mathcal{I} \times \mathcal{J}.$$

The size of the phase-space discretization is defined by $\delta \ = \ \max(\Delta x, \Delta v)$ where Δx and Δv are the maximum of $(\Delta x_i)_{i\in\mathcal{I}}$ and $(\Delta v_j)_{j\in\mathcal{J}}$ respectively. Finally, we set $\Delta t > 0$ the time step, and we define $t^n = n\Delta t$ for all $n \geq 0$.

Discrete Maxwellians. In the case of the BGK operator (1.3), we assume that we are given cell values $\mathcal{M} = (\mathcal{M}_j)_{j\in\mathcal{J}} \in \mathbb{R}^{\mathcal{J}}$ satisfying the following assumptions

$$\left\{ \begin{array}{l} \mathcal{M}_j > 0 \,, \quad \mathcal{M}_j \ = \ \mathcal{M}_{-j+1} \,, \quad \forall j = 1, \dots, L \,; \\ \displaystyle\sum_{j\in\mathcal{J}} \mathcal{M}_j \, \Delta v_j \ = \ 1 \,; \\ 0 \ < \ \underline{m}_2 \ \leq \ m_2^{\Delta v} \ \leq \ \overline{m}_2 \,, \quad m_4^{\Delta v} \ \leq \ \overline{m}_4 \,, \end{array} \right. \tag{3.1}$$

where for $k \in \mathbb{N}$

$$m_k^{\Delta v} \ := \ \sum_{j\in\mathcal{J}} |v_j|^k \, \mathcal{M}_j \, \Delta v_j$$

and $\underline{m}_2$, $\overline{m}_2$, $\overline{m}_4$ are some universal constants.

In the case of the Fokker-Planck operator (1.2), we assume that we are given interface values $(\mathcal{M}^*_{j+1/2})_{j\in\mathcal{J}^*} \in \mathbb{R}^{\mathcal{J}^*}$ such that

$$\left\{ \begin{array}{l} \mathcal{M}^*_{j+1/2} \ = \ \mathcal{M}^*_{-j+1/2} \,, \quad \forall j \in \mathcal{J}^* \,; \\ \mathcal{M}^*_{L+1/2} \ = \ \mathcal{M}^*_{-L+1/2} \ = \ 0 \,; \\ \mathcal{M}_j \ := \ \dfrac{\mathcal{M}^*_{j-1/2} - \mathcal{M}^*_{j+1/2}}{v_j \, \Delta v_j} \ > \ 0 \,, \quad \forall j \in \mathcal{J} \,; \\ \displaystyle\sum_{j\in\mathcal{J}} \mathcal{M}_j \, \Delta v_j \ = \ 1 \,; \\ 0 \ < \ \underline{m}_2 \ \leq \ m_2^{\Delta v} \ \leq \ \overline{m}_2 \,, \quad m_4^{\Delta v} \ \leq \ \overline{m}_4 \,. \end{array} \right. \tag{3.2}$$

These assumptions on the discrete Maxwellians $\mathcal{M}$ are weak and it is easy to build an example.

Example 1. *Define $\mathcal{M}_j := c_{\Delta v} \, \mathcal{M}(v_j)$ in the BGK case and $\mathcal{M}_{j+1/2} := \tilde{c}_{\Delta v} \, \mathcal{M}(v_{j+1/2})$ in the Fokker-Planck case and compute $c_{\Delta v}$ and $\tilde{c}_{\Delta v}$ to normalize the mass of $(\mathcal{M}_j)_j$ to 1. Then, by*

consistency of the piecewise constant approximation (say, in L^1) the last condition in (3.1) *(resp.* (3.2)*) holds in the BGK case (resp. the Fokker-Planck case).*

Remark 3.

- *The second assumption in* (3.2) *is needed for technical reasons, in order to perform discrete integration by parts. However, one can note that for sufficiently large domains this hypothesis is relevant and invisible in practice because of the fast decay of the Gaussian.*
- *In both cases observe that symmetries of both the velocities and the Maxwellians imply*

$$\sum_{j\in\mathcal{J}} \Delta v_j v_j \mathcal{M}_j = 0. \tag{3.3}$$

- *The third assumption in* (3.2) *is the discrete counterpart of* (1.4)*, namely $\partial_v \mathcal{M} = -v\mathcal{M}$.*
- *Let us underline that $m_2^{\Delta v}$ and $m_4^{\Delta v}$ are not consistent with m_2 and m_4 when Δv tends to zero, since we consider the problem in a bounded domain $[-v_*, v_*]$ with zero boundary values for the discrete Maxwellian. We then have an error in $O(\mathcal{M}(v_*))$ for these quantities.*

As in the continuous case, we also introduce the inverse of the Maxwellian

$$\gamma_j \;=\; \frac{1}{\mathcal{M}_j}\,, \quad \forall j \in \mathcal{J}\,.$$

Discrete gradients. Given some discrete microscopic quantity $g = (g_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ defined on control volumes and assuming that some boundary data (in velocity) $(g_{i,L+1})_{i\in\mathcal{I}}$ and $(g_{i,-L})_{i\in\mathcal{I}}$ are given, one may define the discrete gradients $D_v g \in \mathbb{R}^{\mathcal{I}\times\mathcal{J}^*}$ on the dual mesh by

$$(D_v g)_{i,j+1/2} \;=\; \frac{g_{i,j+1} - g_{ij}}{\Delta v_{j+1/2}}\,, \quad \forall j \in \mathcal{J}^*\,, \quad \forall i \in \mathcal{I}.$$

In the numerical analysis of our scheme it will be convenient to use discrete gradients in space too. Given a macroscopic discrete quantity $\rho = (\rho_i)_{i\in\mathcal{I}}$, we define discrete gradients $D_x\rho \in \mathbb{R}^{\mathcal{I}}$ on the (primal) mesh as

$$(D_x\rho)_i \;=\; \frac{\rho_{i+1/2} - \rho_{i-1/2}}{\Delta x_i}\,, \quad \text{where} \quad \rho_{i+1/2} \;=\; \frac{\rho_{i+1} + \rho_i}{2}\,, \quad \forall i \in \mathcal{I}\,.$$

Discrete functional spaces. From there we define for microscopic quantities the discrete weighted L^2 norm

$$\|g\|_{2,\gamma}^2 \;:=\; \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} |g_{ij}|^2\, \gamma_j\, \Delta x_i\, \Delta v_j\,.$$

For macroscopic quantities, we shall need the discrete L^2 norm

$$\|\rho\|_2^2 \;:=\; \sum_{i\in\mathcal{I}} |\rho_i|^2\, \Delta x_i$$

and discrete Sobolev seminorm

$$\|D_x\rho\|_2^2 \;:=\; \sum_{i\in\mathcal{I}} |(D_x\rho)_i|^2\, \Delta x_i\,.$$

Finally, for a discrete Maxwellian satisfying (3.2) one can introduce the Sobolev seminorm in velocity

$$\|D_v g\|_{2,\mathcal{M}^*}^2 \;:=\; \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}^*} \left|(D_v g)_{i,j+1/2}\right|^2\, \mathcal{M}^*_{j+1/2}\, \Delta x_i\, \Delta v_{j+1/2}\,.$$

Observe that since the discrete Maxwellian vanishes at endpoints of the domain, the definition does not involve boundary data $(g_{i,L+1})_{i\in\mathcal{I}}$ and $(g_{i,-L})_{i\in\mathcal{I}}$. One has the following discrete counterparts of the Gaussian and flat Poincaré inequalities. The proof of these results is given in appendix.

Lemma 5 (Discrete Gaussian Poincaré inequality on bounded velocity domain)**.** *Given a discrete Maxwellian on interfaces $(\mathcal{M}^*_{j+1/2})_{j\in\mathcal{J}^*}$ satisfying* (3.2)*, one has for all $f=(f_j)_{j\in\mathcal{J}}\in\mathbb{R}^{\mathcal{J}}$ that*

$$\|f-\rho\,\mathcal{M}\|^2_{2,\gamma} \;\le\; \left\|D_v\left(\frac{f}{\mathcal{M}}\right)\right\|^2_{2,\mathcal{M}^*}, \tag{3.4}$$

where $\mathcal{M}=(\mathcal{M}_j)_{j\in\mathcal{J}}$ is defined in (3.2) *and $\rho=\sum_{j\in\mathcal{J}} f_j\Delta v_j$.*

Lemma 6 (Discrete Poincaré inequality on the torus)**.** *Assume that the number of points N in the space discretization of the torus is odd. Then, there is a constant C_P that does not depend on Δx such that one has for all $\phi=(\phi_i)_{i\in\mathcal{I}}\in\mathbb{R}^{\mathcal{I}}$ satisfying $\sum_{i\in\mathcal{I}}\Delta x_i\phi_i = 0$ that*

$$\|\phi\|_2 \le C_P\,\|D_x\phi\|_2. \tag{3.5}$$

3.2. **Presentation of the schemes.** In this section we present numerical schemes for Equation (1.1) in both the Fokker-Planck and the BGK cases. The schemes are of finite volume type [23], meaning that they are based on an approximation of the fluxes appearing in the integrated version of (1.1) on each cell K_{ij}. In the time variable we choose a backward Euler discretization. First of all, we discretize the initial datum f_0 by

$$f^0_{ij}=\frac{1}{\Delta x_i\,\Delta v_j}\iint_{K_{ij}} f(0,x,v)\,\mathrm{d}x\,\mathrm{d}v,\quad \forall (i,j)\in\mathcal{I}\times\mathcal{J}.$$

Now let us define the schemes corresponding to each collision operator. To lighten the notations, we omit the superscript ε for the discretization of f^ε and its moments in all this section, even if these quantities obviously depend on ε as in the continuous setting.

Fully discrete Fokker-Planck equation. In the case of the Fokker-Planck operator (1.5), the scheme is given as follows. For all $i\in\mathcal{I}$, $j\in\mathcal{J}$, $n\ge 0$,

$$\varepsilon\Delta x_i\Delta v_j(f^{n+1}_{ij}-f^n_{ij})+\Delta t\left(\mathcal{F}^{n+1}_{i+\frac12,j}-\mathcal{F}^{n+1}_{i-\frac12,j}\right)=\frac{\Delta t}{\varepsilon}\left(\mathcal{G}^{n+1}_{i,j+\frac12}-\mathcal{G}^{n+1}_{i,j-\frac12}\right), \tag{3.6}$$

where the numerical fluxes are respectively defined for all $n\ge 0$ by

$$\mathcal{F}^{n+1}_{i+\frac12,j} \;=\; v_j\frac{f^{n+1}_{i+1,j}+f^{n+1}_{ij}}{2}\Delta v_j,\qquad \forall j\in\mathcal{J},\quad \forall i\in\mathcal{I}, \tag{3.7}$$

$$\mathcal{G}^{n+1}_{i,j+\frac12} \;=\; \mathcal{M}^*_{j+\frac12}\left(\frac{f^{n+1}_{i,j+1}}{\mathcal{M}_{j+1}}-\frac{f^{n+1}_{ij}}{\mathcal{M}_j}\right)\frac{1}{\Delta v_{j+\frac12}}\Delta x_i,\qquad \forall j\in\mathcal{J}^*\setminus\{-L,L\},\quad \forall i\in\mathcal{I}, \tag{3.8}$$

$$\mathcal{G}^{n+1}_{i,-L+\frac12} \;=\; \mathcal{G}^{n+1}_{i,L+\frac12} \;=\; 0,\qquad \forall i\in\mathcal{I}. \tag{3.9}$$

Remark 4. *The first equation* (3.7) *is a centered discretization of the free transport. The choice of the velocity fluxes $\mathcal{G}_{i,j+\frac12}$ in* (3.8)*, which is based on the gradient form* (1.5) *of the Fokker-Planck operator*

$$\mathcal{Q}_{FP}(f) \;=\; \partial_v\left(\mathcal{M}\partial_v\left(f/\mathcal{M}\right)\right)$$

is fairly close to the Chang-Cooper [14, 9] *or Scharfetter-Gummel* [34, 50] *approximation. Indeed, if we consider a uniform velocity discretization with $\Delta v_j = \Delta v$ constant, and the discrete Maxwellian defined as in Example 1, the velocity flux is then given by*

$$\mathcal{G}_{i,j+\frac{1}{2}} = \frac{\Delta x_i}{\Delta v}\left(B\left(-v_{j+1}\Delta v\right) f_{i,j+1} - B\left(v_j \Delta v\right) f_{i,j}\right),$$

where B is the Bernoulli function defined by $B(x) = x/(e^x - 1)$ if $x \neq 0$, $B(0) = 1$. The Chang-Cooper and Scharfetter-Gummel schemes are obtained by replacing v_j and v_{j+1} by $v_{j+1/2}$.

Finally, the last equation (3.9) *corresponds to zero flux boundary conditions in velocity.*

Fully discrete BGK equation. In the case of the BGK operator (1.3), the scheme is given as follows. For all $i \in \mathcal{I}$, $j \in \mathcal{J}$, $n \geq 0$,

$$\varepsilon \Delta x_i \Delta v_j (f_{ij}^{n+1} - f_{ij}^n) + \Delta t \left(\mathcal{F}_{i+\frac{1}{2},j}^{n+1} - \mathcal{F}_{i-\frac{1}{2},j}^{n+1}\right) = \frac{\Delta t}{\varepsilon} \Delta x_i \Delta v_j \left(\rho_i^{n+1} \mathcal{M}_j - f_{ij}^{n+1}\right), \tag{3.10}$$

where the numerical flux $\mathcal{F}_{i+\frac{1}{2},j}^{n+1}$ is still defined by (3.7) and for all $i \in \mathcal{I}$ and $n \geq 0$, the discrete macroscopic density is given by

$$\rho_i^n \ := \ \sum_{j \in \mathcal{J}} \Delta v_j f_{ij}^n \,. \tag{3.11}$$

For future use and independently of the choice of collision operator, we also define the other used velocity moments of the discrete distribution, namely

$$J_i^n \ := \ \frac{1}{\varepsilon} \sum_{j \in \mathcal{J}} \Delta v_j \, v_j \, f_{ij}^n \,, \qquad S_i^n \ := \ \sum_{j \in \mathcal{J}} \Delta v_j \left(v_j^2 - m_2^{\Delta v}\right) f_{ij}^n \,, \tag{3.12}$$

for all $i \in \mathcal{I}$ and $n \geq 0$. Both schemes (3.6) and (3.10) clearly satisfy the discrete mass conservation:

$$\sum_{(i,j) \in \mathcal{I} \times \mathcal{J}} \Delta x_i \Delta v_j f_{i,j}^n = \sum_{(i,j) \in \mathcal{I} \times \mathcal{J}} \Delta x_i \Delta v_j f_{i,j}^0 = \mu_f, \quad \forall n \geq 0.$$

This property is obtained by using the zero flux boundary conditions in velocity (3.9) in the Fokker-Planck case, and assumption on the discrete Maxwellian (3.1) in the BGK case.

Existence and uniqueness of a solution to the fully implicit schemes (3.6) and (3.10) will be established in the next section as a by-product of the discrete "entropy" estimate given in Lemma 7.

3.3. **Uniform estimates.** In this section we derive the discrete counterparts of the results of Section 2.1.

Lemma 7 (Discrete "entropy" estimate)**.** *Let us consider a discrete Maxwellian satisfying* (3.2) *(resp.* (3.1)*) and let $(f_{ij}^n)_{i \in \mathcal{I}, j \in \mathcal{J}, n \in \mathbb{N}}$ solve the scheme* (3.6) *(resp.* (3.10)*). Then for every $n \geq 0$,*

$$\frac{\|f^{n+1}\|_{2,\gamma}^2 - \|f^n\|_{2,\gamma}^2}{2\Delta t} \ + \ \frac{1}{\varepsilon^2} \|f^{n+1} - \rho^{n+1} \mathcal{M}\|_{2,\gamma}^2 \ \leq \ 0 \,. \tag{3.13}$$

In particular one has

$$\max\left(\sup_{n \geq 0} \|f^n\|_{2,\gamma}^2, \ \frac{2}{\varepsilon^2} \sum_{n=1}^{\infty} \Delta t \, \|f^n - \rho^n \mathcal{M}\|_{2,\gamma}^2\right) \ \leq \ \|f^0\|_{2,\gamma}^2 \,. \tag{3.14}$$

Since schemes (3.6) and (3.10) are finite dimensional linear systems, we deduce from (3.14) the uniqueness of a solution, and then the existence.

Corollary 1. *The scheme* (3.6) *(resp.* (3.10)*) admits a unique solution $(f_{ij}^n)_{i \in \mathcal{I}, j \in \mathcal{J}, n \in \mathbb{N}}$.*

Proof of Lemma 7. We first consider the Fokker-Planck case. Multiplying (3.6) by $f_{ij}^{n+1}\gamma_j$ and summing over i and j, we get

$$\varepsilon \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} \Delta x_i \Delta v_j (f_{ij}^{n+1} - f_{ij}^n) f_{ij}^{n+1}\gamma_j + \Delta t \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} (\mathcal{F}_{i+\frac{1}{2},j}^{n+1} - \mathcal{F}_{i-\frac{1}{2},j}^{n+1}) f_{ij}^{n+1}\gamma_j = \frac{\Delta t}{\varepsilon} \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} (\mathcal{G}_{i,j+\frac{1}{2}}^{n+1} - \mathcal{G}_{i,j-\frac{1}{2}}^{n+1}) f_{ij}^{n+1}\gamma_j. \tag{3.15}$$

Using that $a(a-b) \geq (a^2-b^2)/2$ for all a, $b \in \mathbb{R}$, the first term of (3.15) can be bounded from below by

$$\frac{\varepsilon}{2} \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} \Delta x_i \Delta v_j \left((f_{ij}^{n+1})^2 - (f_{ij}^n)^2 \right) \gamma_j = \frac{\varepsilon}{2} (\|f^{n+1}\|_{2,\gamma}^2 - \|f^n\|_{2,\gamma}^2).$$

Then, using the definition (3.7) of the numerical fluxes, the second term of (3.15) gives

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} (\mathcal{F}_{i+\frac{1}{2},j}^{n+1} - \mathcal{F}_{i-\frac{1}{2},j}^{n+1}) f_{ij}^{n+1}\gamma_j &= \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} \Delta v_j v_j \gamma_j \left(\frac{f_{i+1,j}^{n+1} + f_{ij}^{n+1}}{2} - \frac{f_{ij}^{n+1} + f_{i-1,j}^{n+1}}{2} \right) f_{ij}^{n+1} \\
&= \sum_{j\in\mathcal{J}} \Delta v_j v_j \gamma_j \sum_{i\in\mathcal{I}} \frac{f_{i+1,j}^{n+1} f_{ij}^{n+1} - f_{ij}^{n+1} f_{i-1,j}^{n+1}}{2} = 0.
\end{aligned}$$

Finally, we perform a discrete integration by parts in velocity on the third term of (3.15) using the zero flux boundary conditions, to obtain

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} (\mathcal{G}_{i,j+\frac{1}{2}}^{n+1} - \mathcal{G}_{i,j-\frac{1}{2}}^{n+1}) f_{ij}^{n+1}\gamma_j &= - \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}^*} \mathcal{G}_{i,j+\frac{1}{2}}^{n+1} (f_{i,j+1}^{n+1}\gamma_{j+1} - f_{ij}^{n+1}\gamma_j) \\
&= - \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}^*} \frac{\Delta x_i}{\Delta v_{j+\frac{1}{2}}} \mathcal{M}_{j+\frac{1}{2}}^* (f_{i,j+1}^{n+1}\gamma_{j+1} - f_{ij}^{n+1}\gamma_j)^2 \\
&= -\|D_v(f/\mathcal{M})\|_{2,\mathcal{M}^*}^2 \\
&\leq -\|f^{n+1} - \rho^{n+1}\mathcal{M}\|_{2,\gamma}^2,
\end{aligned}$$

thanks to the discrete Gaussian Poincaré inequality. This concludes the proof in the Fokker-Planck case.

For the BGK operator the left hand side is exactly the same as in (3.15). Using the definition (3.11) of ρ_i^{n+1} and the unit mass assumption on the discrete Mawxellian in (3.1), we obtain for the right hand side

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} \Delta x_i \Delta v_j (\rho_i^{n+1}\mathcal{M}_j - f_{ij}^{n+1}) f_{ij}^{n+1}\gamma_j &= - \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} \Delta x_i \Delta v_j (f_{ij}^{n+1} - \rho_i^{n+1}\mathcal{M}_j)^2 \gamma_j \\
&= -\|f^{n+1} - \rho^{n+1}\mathcal{M}\|_{2,\gamma}^2,
\end{aligned}$$

which concludes the proof. □

Lemma 8 (Discrete moments estimates)**.** *Under the hypotheses of Lemma 7, the moments* $(\rho_i^n)_{i\in\mathcal{I}}$, $(J_i^n)_{i\in\mathcal{I}}$ *and* $(S_i^n)_{i\in\mathcal{I}}$ *satisfy the following estimates, for all* $n \in \mathbb{N}$*:*

$$\|\rho^n\|_2 \leq \|f^n\|_{2,\gamma}, \tag{3.16}$$

$$\varepsilon \|J^n\|_2 \leq (m_2^{\Delta v})^{1/2} \|f^n\|_{2,\gamma}, \tag{3.17}$$

$$\varepsilon \|J^n\|_2 \leq (m_2^{\Delta v})^{1/2} \|f^n - \rho^n \mathcal{M}\|_{2,\gamma}, \tag{3.18}$$

$$\|S^n\|_2 \leq (m_4^{\Delta v} - (m_2^{\Delta v})^2)^{1/2} \|f^n - \rho^n \mathcal{M}\|_{2,\gamma}. \tag{3.19}$$

In particular, there is a constant $C > 0$ depending only on the uniform bounds on $m_2^{\Delta v}$ and $m_4^{\Delta v}$ from hypotheses (3.1) *and* (3.2) *such that*

$$\max\left(\sup_{n\in\mathbb{N}}\|\rho^n\|_2,\, \varepsilon\sup_{n\in\mathbb{N}}\|J^n\|_2,\, \left(\sum_{n=0}^{\infty}\Delta t\,\|J^n\|_2^2\right)^{1/2},\, \frac{1}{\varepsilon}\left(\sum_{n=0}^{\infty}\Delta t\,\|S^n\|_2^2\right)^{1/2}\right) \leq C\,\|f^0\|_{2,\gamma}\,. \tag{3.20}$$

Proof. The proof is almost identical to that of Lemma 2. Indeed, just observe that the moments rewrite as follows. First one has

$$\rho_i^n \;=\; \sum_{j\in\mathcal{J}}\Delta v_j\,(f_{ij}^n\gamma_j^{1/2})\mathcal{M}_j^{1/2}\,.$$

Then, thanks to the symmetry of the discrete Maxwellian and anti-symmetry of discrete velocities yielding (3.3), one has

$$J_i^n \;=\; \frac{1}{\varepsilon}\sum_{j\in\mathcal{J}}\Delta v_j\,(f_{ij}^n\gamma_j^{1/2})\,(v_j\,\mathcal{M}_j^{1/2}) \;=\; \frac{1}{\varepsilon}\sum_{j\in\mathcal{J}}\Delta v_j\,(f_{ij}^n-\rho_i^n\,\mathcal{M}_j)\gamma_j^{1/2}\,(v_j\,\mathcal{M}_j^{1/2})\,.$$

Finally, observe that by definition of $m_2^{\Delta v}$ one has

$$S_i^n \;=\; \sum_{j\in\mathcal{J}}\Delta v_j\,(f_{ij}^n-\rho_i^n\,\mathcal{M}_j)\gamma_j^{1/2}\,(v_j^2-m_2^{\Delta v})\,\mathcal{M}_j^{1/2}\,.$$

Then one can take the norm of each expression and use Cauchy-Schwarz inequalities to get the desired estimates. Estimate (3.20) is then a direct consequence of (3.14). □

3.4. **Asymptotic-preserving property.** For macroscopic quantities we recall that approximate values at the interfaces $x_{i+\frac12}$, $i\in\mathcal{I}$ are defined by the average

$$Q_{i+\frac12}=\frac{Q_i+Q_{i+1}}{2},\quad X=\rho,\,J,\,S.$$

Lemma 9 (Moments equations)**.** *Let us consider a discrete Maxwellian satisfying* (3.2) *(resp.* (3.1)*) and a solution to the scheme* (3.6) *(resp.* (3.10)*). Then the discrete moments satisfy the following equations. For all $i\in\mathcal{I}$, $n\geq 0$,*

$$\Delta x_i\,(\rho_i^{n+1}-\rho_i^n)\;+\;\Delta t\,(J_{i+\frac12}^{n+1}-J_{i-\frac12}^{n+1}) \;=\; 0, \tag{3.21}$$

$$\varepsilon^2\,\Delta x_i\,(J_i^{n+1}-J_i^n)\;+\;\Delta t\,(S_{i+\frac12}^{n+1}-S_{i-\frac12}^{n+1})\;+\;\Delta t\,m_2^{\Delta v}\,(\rho_{i+\frac12}^{n+1}-\rho_{i-\frac12}^{n+1}) \;=\; -\,\Delta t\,\Delta x_i\,J_i^{n+1}. \tag{3.22}$$

Proof. Equation (3.21) can be easily obtained by taking (3.6) (resp. (3.10)) and summing over $j\in\mathcal{J}$. Thanks to the zero flux boundary conditions in the Fokker-Planck case, and the normalization assumption of the Maxwellian in the BGK case, one easily sees that the contribution of the collision operators vanishes.

To obtain the second equation, we multiply (3.6) (resp. (3.10)) by v_j and sum over j. In the Fokker-Planck case, it yields

$$\varepsilon^2\Delta x_i(J_i^{n+1}-J_i^n)+\Delta t(S_{i+\frac12}^{n+1}-S_{i-\frac12}^{n+1})+\Delta t m_2^{\Delta v}(\rho_{i+\frac12}^{n+1}-\rho_{i-\frac12}^{n+1})= \frac{\Delta t}{\varepsilon}\sum_{j\in\mathcal{J}}v_j\left(\mathcal{G}_{i,j+\frac12}^{n+1}-\mathcal{G}_{i,j-\frac12}^{n+1}\right).$$

Performing a discrete integration by parts and using zero flux boundary conditions, we have

$$\frac{\Delta t}{\varepsilon}\sum_{j\in\mathcal{J}} v_j\left(\mathcal{G}^{n+1}_{i,j+\frac12}-\mathcal{G}^{n+1}_{i,j-\frac12}\right) = -\frac{\Delta t}{\varepsilon}\sum_{j\in\mathcal{J}^*}(v_{j+1}-v_j)\mathcal{G}^{n+1}_{i,j+\frac12}$$
$$= -\frac{\Delta t}{\varepsilon}\,\Delta x_i \sum_{j\in\mathcal{J}^*\setminus\{-L,L\}} \mathcal{M}^*_{j+\frac12}\left(\frac{f^{n+1}_{i,j+1}}{\mathcal{M}_{j+1}}-\frac{f^{n+1}_{ij}}{\mathcal{M}_j}\right).$$

Now, performing again a discrete integration by parts and using that the discrete Maxwellian vanishes at endpoints of the velocity domain (see assumption (3.2)), we obtain

$$\frac{\Delta t}{\varepsilon}\sum_{j\in\mathcal{J}} v_j\left(\mathcal{G}^{n+1}_{i,j+\frac12}-\mathcal{G}^{n+1}_{i,j-\frac12}\right) = \frac{\Delta t}{\varepsilon}\,\Delta x_i\sum_{j\in\mathcal{J}}(\mathcal{M}^*_{j+\frac12}-\mathcal{M}^*_{j-\frac12})\frac{f^{n+1}_{ij}}{\mathcal{M}_j}$$
$$= -\frac{\Delta t}{\varepsilon}\,\Delta x_i\sum_{j\in\mathcal{J}}\Delta v_j v_j f^{n+1}_{ij},$$

which finally yields the result using definition (3.12) of J^{n+1}_i.

In the BGK case, we also multiply (3.10) by v_j and sum over j to obtain

$$\varepsilon^2\Delta x_i(J^{n+1}_i-J^n_i)+\Delta t(S^{n+1}_{i+\frac12}-S^{n+1}_{i-\frac12})+\Delta t m_2^{\Delta v}(\rho^{n+1}_{i+\frac12}-\rho^{n+1}_{i-\frac12}) =$$
$$\frac{\Delta t}{\varepsilon}\Delta x_i\sum_{j\in\mathcal{J}}\Delta v_j(\rho^{n+1}_i\mathcal{M}_j-f^{n+1}_{ij})v_j.$$

Using (3.3), the right hand side can be written as $-\Delta t\Delta x_i J^{n+1}_i$, which concludes the proof. □

Theorem 1. *(Asymptotic-preserving property) Let us consider a discrete Maxwellian satisfying* (3.2) *(resp.* (3.1)*) and let $f^n_\varepsilon=(f^n_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ for $n\in\mathbb{N}$ be the solution of* (3.6) *(resp.* (3.10)*) with initial data $(f^0_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$. Then there is $\rho^n=(\rho^n_i)_{i\in\mathcal{I}}$ for all $n\geq 0$ such that when $\varepsilon\to 0$ one has*

$$f^n_\varepsilon \longrightarrow \rho^n\mathcal{M} \quad in \quad \mathbb{R}^{\mathcal{I}\times\mathcal{J}}, \quad for\ all \quad n\geq 1$$

and the limit satisfies the following finite difference scheme for the heat equation

$$\Delta x_i\frac{\rho^{n+1}_i-\rho^n_i}{\Delta t} = \frac{m_2^{\Delta v}}{2}\left((D_x\rho^{n+1})_{i+1}-(D_x\rho^{n+1})_{i-1}\right), \quad \forall i\in\mathcal{I} \tag{3.23}$$

with initial data $\rho^0_i=\sum_{j\in\mathcal{J}}\Delta v_j f^0_{ij}$.

Proof. Thanks to estimate (3.20) on the discrete macroscopic density, there is a subsequence $(\varepsilon_k)_k$ such that $\rho^n_{\varepsilon_k}\to\rho^n$ for all $n\geq 0$. Thanks to estimate (3.14) the sequence $(f^n_{\varepsilon_k})_k\in(\mathbb{R}^{\mathcal{I}\times\mathcal{J}})^{\mathbb{N}}$ converges to $(\rho^n_i\,\mathcal{M}_j)_{i\in\mathcal{I},j\in\mathcal{J}}$ for each $n\geq 1$. To identify the limit scheme, we combine (3.21) and (3.22) to get that for $\varepsilon>0$ one has for all $i\in\mathcal{I}$ and $n\in\mathbb{N}$ that

$$\Delta x_i\,(\rho^{n+1}_i-\rho^n_i) = m_2^{\Delta v}\,\Delta t\,\frac12\left((D_x\rho^{n+1})_{i+1}-(D_x\rho^{n+1})_{i-1}\right)$$
$$+\ \Delta t\,\frac12\left((D_xS^{n+1})_{i+1}-(D_xS^{n+1})_{i-1}\right)+\varepsilon^2\,(J^{n+1}_{i+\frac12}-J^{n+1}_{i-\frac12})-\varepsilon^2\,(J^n_{i+\frac12}-J^n_{i-\frac12}).$$

Then thanks to estimate (3.20) it is straightforward to show that the terms of the second line converge to 0. Finally observe that the limit scheme has a unique solution. Indeed, it consists in inverting a matrix that is strictly diagonally dominant and thus that is non singular. Therefore the whole sequence converges. □

Remark 5. *Observe that the scheme for the heat equation keeps a trace of the velocity discretization in the diffusion coefficient $m_2^{\Delta v}$.*

3.5. **Numerical hypocoercivity.** We shall now adapt to the discrete framework the hypocoercivity method proposed by Dolbeault, Mouhot and Schmeiser in [21] and summarized in Section 2.3. The main result of the section is the following.

Theorem 2. *Let us consider a discrete Maxwellian satisfying* (3.2) *(resp.* (3.1)*) and assume that the number of points N in the space discretization is odd. Then there are constants $C \geq 1$ and $\beta > 0$ such that for all $\varepsilon \in (0,1)$, all $\Delta t \leq \Delta t_{\max}$ and all initial data $(f_{ij}^0)_{i\in\mathcal{I},j\in\mathcal{J}}$, the solution $(f_{ij}^n)_{i\in\mathcal{I},j\in\mathcal{J},n\in\mathbb{N}}$ of* (3.6) *(resp.* (3.10)*) satisfies*

$$\|f^n - \mu_f\mathcal{M}\|_{2,\gamma} \;\leq\; C\,\|f^0 - \mu_f\mathcal{M}\|_{2,\gamma}\,e^{-\frac{\beta}{2}\,t^n}\,. \tag{3.24}$$

Moreover, the constants C and β do not depend on the size of the discretization δ, and $\Delta t_{\max} > 0$ can be chosen arbitrarily.

Once again we prove the theorem assuming that the initial data has zero mean

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} \Delta x_i\,\Delta v_j\,f_{ij}^0 \;=\; 0\,.$$

In the general case one can just replace f_{ij}^0 by $f_{ij}^0 - \mu_f\mathcal{M}_j$, where $\mu_f = \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}}\Delta x_i\Delta v_j f_{ij}^0$, and correspondingly f_{ij}^n and ρ_i^n by respectively $f_{ij}^n - \mu_f\mathcal{M}_j$ and $\rho_i^n - \mu_f$ in the following arguments.

For all $n \geq 1$, let us define the *discrete modified entropy functional* which reads

$$\mathbb{H}(f^n) \;:=\; \frac{1}{2}\|f^n\|_{2,\gamma}^2 + \eta\varepsilon^2\sum_{i\in\mathcal{I}}\Delta x_i J_i^n (D_x\phi)_i^n + \frac{\eta\varepsilon^2}{2}\sum_{i\in\mathcal{I}}\Delta x_i\frac{\left((D_x\phi)_i^n - (D_x\phi)_i^{n-1}\right)^2}{\Delta t}, \tag{3.25}$$

where $\eta > 0$ will be determined later and $(\phi_i^n)_{i\in\mathcal{I}}$ is the solution of the following discrete Poisson equation

$$-\frac{(D_x\phi)_{i+1}^n - (D_x\phi)_{i-1}^n}{2} \;=\; \Delta x_i\rho_i^n, \quad \forall i\in\mathcal{I}, \tag{3.26}$$

$$\sum_{i\in\mathcal{I}}\Delta x_i\phi_i^n \;=\; 0. \tag{3.27}$$

Remark 6. *Observe that the discrete version of the modified entropy has an additional third term compared to* (2.9)*. It is of order $O(\Delta t)$ and thus consistent with* 0*, so that taking limits in the discretization parameters, we recover (at least formally at this stage) the continuous modified entropy* (2.9)*. It does not perturb the adaptation of the strategy of Section 2.3.*

Let us first justify existence and unicity of $(\phi_i)_{i\in\mathcal{I}}$ satisfying (3.26)-(3.27). System (3.26) can be written under the matrix form $A\phi^n = \rho^n$, where the rank of the $N\times N$ matrix A is either $N-1$ if N is odd, or $N-2$ if N is even. Observe that A is symmetric with respect to the inner product $\langle\phi,\psi\rangle = \sum_{i\in\mathcal{I}}\phi_i\,\psi_i\,\Delta x_i$. In the odd case, the condition (3.27) corresponds exactly to $(\phi_i)_{i\in\mathcal{I}}$ being orthogonal (for the latter inner product) to the kernel of A spanned by $(1,\dots,1)\in\mathbb{R}^{\mathcal{I}}$. Hence provided that N is odd, which is assumed in Theorem 2, the system (3.26) supplemented with zero mean value condition (3.27) admits a unique solution.

Let us now derive the discrete estimates on $(\phi_i)_{i\in\mathcal{I}}$.

Lemma 10. *Under the assumptions of Theorem 2, one has for all $n\in\mathbb{N}$*

$$\|D_x\phi^n\|_2 \;\leq\; C_P\,\|\rho^n\|_2\,, \tag{3.28}$$

$$\left\|\frac{(D_x\phi)_i^{n+1} - (D_x\phi)_i^n}{\Delta t}\right\|_2 \;\leq\; \|J_i^{n+1}\|_2\,, \tag{3.29}$$

where C_P is the discrete Poincaré constant of Lemma 6.

Proof. Multiplying the discrete Poisson equation (3.26) by ϕ_i^n, summing on $i \in \mathcal{I}$ and performing a discrete integration by parts, we obtain

$$-\sum_{i\in\mathcal{I}} \Delta x_i \left((D_x\phi)_i^n\right)^2 = -\sum_{i\in\mathcal{I}} \Delta x_i \rho_i^n \phi_i^n,$$

which yields thanks to Cauchy-Schwarz inequality

$$\sum_{i\in\mathcal{I}} \Delta x_i \left((D_x\phi)_i^n\right)^2 \leq \left(\sum_{i\in\mathcal{I}} \Delta x_i (\rho_i^n)^2\right)^{\frac{1}{2}} \left(\sum_{i\in\mathcal{I}} \Delta x_i (\phi_i^n)^2\right)^{\frac{1}{2}}.$$

Using the Poincaré inequality (3.5) we get the first estimate.

Now, considering scheme (3.26) at time n and $n+1$, we can write

$$-\frac{1}{2}\left[\left((D_x\phi)_{i+1}^{n+1} - (D_x\phi)_{i-1}^{n+1}\right) - \left((D_x\phi)_{i+1}^{n} - (D_x\phi)_{i-1}^{n}\right)\right] \;=\; \Delta x_i(\rho_i^{n+1} - \rho_i^n).$$

Multiplying this equality by $\phi_i^{n+1} - \phi_i^n$, summing on $i \in \mathcal{I}$ and integrating by parts, we get

$$\sum_{i\in\mathcal{I}} \Delta x_i \left[(D_x\phi)_i^{n+1} - (D_x\phi)_i^n\right]^2 = \sum_{i\in\mathcal{I}} \Delta x_i (\rho_i^{n+1} - \rho_i^n)(\phi_i^{n+1} - \phi_i^n).$$

Using the scheme (3.21), performing a discrete integration by parts and applying Cauchy-Schwarz inequality, we obtain:

$$\begin{aligned}
\sum_{i\in\mathcal{I}} \Delta x_i \left[(D_x\phi)_i^{n+1} - (D_x\phi)_i^n\right]^2 &= -\Delta t \sum_{i\in\mathcal{I}} \frac{J_{i+1}^{n+1} - J_{i-1}^{n+1}}{2}(\phi_i^{n+1} - \phi_i^n) \\
&= \Delta t \sum_{i\in\mathcal{I}} \Delta x_i J_i^{n+1}\left[(D_x\phi)_i^{n+1} - (D_x\phi)_i^n\right] \\
&\leq \Delta t \left(\sum_{i\in\mathcal{I}} \Delta x_i (J_i^{n+1})^2\right)^{\frac{1}{2}} \left(\sum_{i\in\mathcal{I}} \Delta x_i \left[(D_x\phi)_i^{n+1} - (D_x\phi)_i^n\right]^2\right)^{\frac{1}{2}},
\end{aligned}$$

from which we deduce (3.29). □

Now we show that for η small enough the modified entropy functional (3.25) is an equivalent $\|\cdot\|_{2,\gamma}$ norm uniformly for small ε and mesh size.

Lemma 11. *Under the assumptions of Theorem 2, and assuming that $\varepsilon \leq 1$ and $\Delta t \leq \Delta t_{max}$, one has for all $n \geq 1$*

(3.30)
$$\left(\frac{1}{2} - \eta\,(m_2^{\Delta v})^{1/2}\right)\|f^n\|_{2,\gamma}^2 \;\leq\; \mathbb{H}(f^n) \;\leq\; \left(\frac{1}{2} + \eta\,(m_2^{\Delta v})^{1/2} C_P + \frac{\eta}{2} m_2^{\Delta v}\Delta t_{max}\right)\|f^n\|_{2,\gamma}^2.$$

Proof. On the one hand, using the Cauchy-Schwarz inequality, we have

$$\eta\varepsilon^2 \left|\sum_{i\in\mathcal{I}} \Delta x_i J_i^n (D_x\phi)_i^n\right| \leq \eta\varepsilon^2 \left(\sum_{i\in\mathcal{I}} \Delta x_i (J_i^n)^2\right)^{\frac{1}{2}} \left(\sum_{i\in\mathcal{I}} \Delta x_i \left((D_x\phi)_i^n\right)^2\right)^{\frac{1}{2}},$$

so that with (3.17), (3.28) and (3.16) one has

$$\eta\varepsilon^2 \left|\sum_{i\in\mathcal{I}} \Delta x_i J_i^n (D_x\phi)_i^n\right| \;\leq\; \eta\,(m_2^{\Delta v})^{1/2}\, C_P\, \varepsilon\, \|f^n\|_{2,\gamma}^2.$$

On the other hand the last term of $\mathbb{H}(f^n)$ may be estimated using (3.29) and (3.17)

$$\frac{\eta\varepsilon^2}{2}\sum_{i\in\mathcal{I}}\Delta x_i\frac{\left((D_x\phi)_i^n-(D_x\phi)_i^{n-1}\right)^2}{\Delta t} \;\leq\; \frac{\eta}{2}\,m_2^{\Delta v}\,\Delta t\,\|f^{n+1}\|_{2,\gamma}^2.$$

We finally establish (3.30) using that $\varepsilon\leq 1$ and $\Delta t\leq\Delta t_{max}$. □

Proposition 3. *Under the assumptions of Theorem 2, there is $\eta_2>0$ such that for all $\varepsilon\leq 1$, $\Delta t\leq\Delta t_{max}$ and $\eta\leq\eta_2$,*

$$\frac{\mathbb{H}(f^{n+1})-\mathbb{H}(f^n)}{\Delta t}+K(\eta)\left(\|f^{n+1}-\rho^{n+1}\mathcal{M}\|_{2,\gamma}^2+\|\rho^{n+1}\|_2^2\right)\;\leq\;0\,,\qquad\forall n\geq 1\,,$$

with $K(\eta)=\frac{1}{2}\min\left(1-\eta\overline{m}_2,\eta\underline{m}_2\right)$.

Proof. Let us estimate the discrete time derivative of $\mathbb{H}$. For all $n\geq 1$, we have

$$\mathbb{H}(f^{n+1})-\mathbb{H}(f^n)\;=\;\frac{1}{2}(\|f^{n+1}\|_{2,\gamma}^2-\|f^n\|_{2,\gamma}^2)+\eta\varepsilon^2T_1^n+\eta\varepsilon^2T_2^n, \tag{3.31}$$

with

$$T_1^n=\sum_{i\in\mathcal{I}}\Delta x_i\left(J_i^{n+1}(D_x\phi)_i^{n+1}-J_i^n(D_x\phi)_i^n\right),$$
$$T_2^n=\frac{1}{2}\sum_{i\in\mathcal{I}}\frac{\Delta x_i}{\Delta t}\left(\left((D_x\phi)_i^{n+1}-(D_x\phi)_i^n\right)^2-\left((D_x\phi)_i^n-(D_x\phi)_i^{n-1}\right)^2\right).$$

Concerning the first term, we proved in Lemma 7 that

$$\frac{1}{2}(\|f^{n+1}\|_{2,\gamma}^2-\|f^n\|_{2,\gamma}^2)\leq-\frac{\Delta t}{\varepsilon^2}\|f^{n+1}-\rho^{n+1}\mathcal{M}\|_{2,\gamma}^2. \tag{3.32}$$

Then, we can write:

$$T_1^n=\sum_{i\in\mathcal{I}}\Delta x_i(J_i^{n+1}-J_i^n)(D_x\phi)_i^{n+1}+\sum_{i\in\mathcal{I}}\Delta x_iJ_i^{n+1}\left((D_x\phi)_i^{n+1}-(D_x\phi)_i^n\right)+R^n, \tag{3.33}$$

where

$$R^n:=\sum_{i\in\mathcal{I}}\Delta x_i\left(J_i^n(D_x\phi)_i^{n+1}-J_i^{n+1}(D_x\phi)_i^{n+1}+J_i^{n+1}(D_x\phi)_i^n-J_i^n(D_x\phi)_i^n\right).$$

We remark that the two first sums in the right hand side of (3.33) correspond to a discrete version of $(\langle\partial_tj^\varepsilon,\partial_x\phi^\varepsilon\rangle+\langle j^\varepsilon,\partial_t\partial_x\phi^\varepsilon\rangle)\Delta t$.

Let us now study the remainder term R^n. We have

$$R^n=\sum_{i\in\mathcal{I}}\Delta x_i(J_i^n-J_i^{n+1})\left((D_x\phi)_i^{n+1}-(D_x\phi)_i^n\right),$$

which after a discrete integration by parts leads to

$$R^n=\sum_{i\in\mathcal{I}}\Delta x_i\left((D_xJ)_i^{n+1}-(D_xJ)_i^n\right)(\phi_i^{n+1}-\phi_i^n).$$

Using scheme (3.21), we get

$$R^n=-\frac{1}{\Delta t}\sum_{i\in\mathcal{I}}\Delta x_i(\rho_i^{n+1}-2\rho_i^n+\rho_i^{n-1})(\phi_i^{n+1}-\phi_i^n),$$

and then using scheme (3.26) followed by a discrete integration by parts, we obtain:

$$\begin{aligned} R^n &= \frac{1}{2\Delta t}\sum_{i\in\mathcal{I}}\big[\big((D_x\phi)_{i+1}^{n+1}-(D_x\phi)_{i-1}^{n+1}\big)-2\big((D_x\phi)_{i+1}^{n}-(D_x\phi)_{i-1}^{n}\big) \\ &\qquad +\big((D_x\phi)_{i+1}^{n-1}-(D_x\phi)_{i-1}^{n-1}\big)\big](\phi_i^{n+1}-\phi_i^n) \\ &= -\frac{1}{\Delta t}\sum_{i\in\mathcal{I}}\Delta x_i\big((D_x\phi)_i^{n+1}-2(D_x\phi)_i^n+(D_x\phi)_i^{n-1}\big)\big((D_x\phi)_i^{n+1}-(D_x\phi)_i^n\big)\,. \end{aligned}$$

Using that $-a(a-b)+(a^2-b^2)/2 = -(a-b)^2/2 \le 0$ with $a = (D_x\phi)_i^{n+1}-(D_x\phi)_i^n$ and $b=(D_x\phi)_i^n-(D_x\phi)_i^{n-1}$, we finally get

$$R^n + T_2^n \le 0\,. \tag{3.34}$$

Then gathering (3.32) and (3.34) in (3.31) results in

$$\mathbb{H}(f^{n+1})-\mathbb{H}(f^n) \le -\frac{\Delta t}{\varepsilon^2}\|f^{n+1}-\rho^{n+1}\mathcal{M}\|_{2,\gamma}^2 + T_3^n, \tag{3.35}$$

with

$$T_3^n := \eta\varepsilon^2\sum_{i\in\mathcal{I}}\Delta x_i(J_i^{n+1}-J_i^n)(D_x\phi)_i^{n+1}+\eta\varepsilon^2\sum_{i\in\mathcal{I}}\Delta x_i J_i^{n+1}\big((D_x\phi)_i^{n+1}-(D_x\phi)_i^n\big)\,.$$

It remains now to estimate this term T_3^n. We proceed as in the continuous framework. Using scheme (3.22) in the first sum of the right hand side, we write:

$$T_3^n = T_{31}^n + T_{32}^n + T_{33}^n + T_{34}^n,$$

where

$$\begin{aligned} T_{31}^n &:= -\eta\Delta t\sum_{i\in\mathcal{I}}\Delta x_i(D_xS)_i^{n+1}(D_x\phi)_i^{n+1}, \\ T_{32}^n &:= -\eta\Delta t\sum_{i\in\mathcal{I}}\Delta x_i m_2^{\Delta v}(D_x\rho)_i^{n+1}(D_x\phi)_i^{n+1}, \\ T_{33}^n &:= -\eta\Delta t\sum_{i\in\mathcal{I}}\Delta x_i J_i^{n+1}(D_x\phi)_i^{n+1}, \\ T_{34}^n &:= \eta\varepsilon^2\sum_{i\in\mathcal{I}}\Delta x_i J_i^{n+1}\big((D_x\phi)_i^{n+1}-(D_x\phi)_i^n\big)\,. \end{aligned}$$

Applying a discrete integration by parts and using scheme (3.26), we have

$$T_{31}^n = \eta\Delta t\sum_{i\in\mathcal{I}}S_i^{n+1}\frac{(D_x\phi)_{i+1}^{n+1}-(D_x\phi)_{i-1}^{n+1}}{2} = -\eta\Delta t\sum_{i\in\mathcal{I}}\Delta x_i S_i^{n+1}\rho_i^{n+1}\,.$$

Then the Cauchy-Schwarz inequality and (3.19) yields

$$|T_{31}^n| \le \eta\Delta t\|S^{n+1}\|_2\|\rho^{n+1}\|_2 \le \eta\Delta t\varepsilon\sqrt{\overline{m}_4-\underline{m}_2^2}\,\frac{\|f^{n+1}-\rho^{n+1}\mathcal{M}\|_{2,\gamma}}{\varepsilon}\|\rho^{n+1}\|_2\,. \tag{3.36}$$

Performing a discrete integration by parts and using again scheme (3.26), we get

$$T_{32}^n \;=\; -\eta\Delta t m_2^{\Delta v}\sum_{i\in\mathcal{I}}\Delta x_i(\rho_i^{n+1})^2 \;=\; -\eta\,\Delta t\,m_2^{\Delta v}\,\|\rho^{n+1}\|_2^2\,. \tag{3.37}$$

Then, applying Cauchy-Schwarz inequality followed by (3.18) and (3.28) we get

$$|T_{33}^n| \le \eta\,\Delta t\,\|J^{n+1}\|_2\,\|D_x\phi^{n+1}\|_2 \le \eta\,\Delta t\,C_P\,\sqrt{\overline{m}_2}\,\frac{\|f^{n+1}-\rho^{n+1}\mathcal{M}\|_{2,\gamma}}{\varepsilon}\|\rho^{n+1}\|_2\,. \tag{3.38}$$

The last remainder term is estimated again using Cauchy-Schwarz inequality followed by (3.18) and (3.29), yielding

$$|T_{34}^n| \leq \eta\varepsilon^2 \, \|J^{n+1}\|_2 \, \|D_x\phi^{n+1} - D_x\phi^n\|_2 \leq \eta\Delta t\overline{m}_2 \|f^{n+1} - \rho^{n+1}\mathcal{M}\|_{2,\gamma}^2. \tag{3.39}$$

Gathering (3.36), (3.37), (3.38) and (3.39) in (3.35) then gives:

$$\begin{aligned}\mathbb{H}(f^{n+1}) - \mathbb{H}(f^n) + \Delta t(\frac{1}{\varepsilon^2} - \eta\overline{m}_2)\|f^{n+1} - \rho^{n+1}\mathcal{M}\|_{2,\gamma}^2 + \Delta t\eta\underline{m}_2\|\rho^{n+1}\|_2^2 \\ \leq \Delta t\eta(\sqrt{\overline{m}_4 - \underline{m}_2^2} + \frac{C_P\sqrt{\overline{m}_2}}{\varepsilon})\|f^{n+1} - \rho^{n+1}\mathcal{M}\|_{2,\gamma}\|\rho^{n+1}\|_2.\end{aligned}$$

It follows that for any $\eta \in (0, \eta_2)$ with

$$\eta_2 := \frac{\underline{m}_2}{(\sqrt{\overline{m}_4 - \underline{m}_2^2} + C_P\sqrt{\overline{m}_2})^2 + \underline{m}_2\overline{m}_2},$$

and $\varepsilon \in (0,1)$, one has

$$\mathbb{H}(f^{n+1}) - \mathbb{H}(f^n) + \Delta t K(\eta)\left(\|f^{n+1} - \rho^{n+1}\mathcal{M}\|_{2,\gamma}^2 + \|\rho^{n+1}\|_2^2\right) \leq 0,$$

with

$$K(\eta) = \frac{1}{2}\min\left(1 - \eta\overline{m}_2, \eta\underline{m}_2\right).$$

□

Proof of Theorem 2. To conclude the proof we observe that $f^{n+1} - \rho^{n+1}\mathcal{M}$ and $\rho^{n+1}\mathcal{M}$ are orthogonal with respect to $\|\cdot\|_{2,\gamma}$ and thus with (3.30) one obtains

$$\mathbb{H}(f^{n+1}) - \mathbb{H}(f^n) + \Delta t\kappa(\eta)\,\mathbb{H}(f^{n+1}) \leq 0 \quad \forall n \geq 1,$$

and $\kappa(\eta) = 2K(\eta)/K_2(\eta)$ where $K_2(\eta) = 1 + 2\eta\sqrt{\overline{m}_2}C_P + \eta\overline{m}_2\Delta t_{\max}$. It implies

$$\mathbb{H}(f^n) \;\leq\; \mathbb{H}(f^1)\,(1 + \Delta t\,\kappa(\eta))^{-(n-1)} \;= \mathbb{H}(f^1)\exp(-t^{n-1}(\Delta t)^{-1}\log(1 + \Delta t\,\kappa(\eta)))\,.$$

Since $s \mapsto \log(1 + s\kappa(\eta))/s$ is a non increasing function on $]0, +\infty[$ and $\Delta t \leq \Delta t_{max}$, one has

$$\mathbb{H}(f^{n+1}) \leq e^{\beta\Delta t_{max}}\mathbb{H}(f^1)e^{-\beta t^n} \quad \forall n \geq 1,$$

with maximal uniform rate

$$\beta = \frac{1}{\Delta t_{max}}\log(1 + \Delta t_{max}\kappa(\eta))\,.$$

Finally

$$\mathbb{H}(f^1) \leq \frac{K_2(\eta)}{2}\|f^1\|_{2,\gamma}^2 \leq \frac{K_2(\eta)}{2}\|f^0\|_{2,\gamma}^2\,.$$

Provided that $0 < \eta < \min(\eta_1, \eta_2)$, with $\eta_1 = 1/(2\underline{m}_2^{1/2})$ to ensure that the constant at the left hand side of (3.30) is positive, one finally obtains (3.24), which concludes the proof of Theorem 2. □

4. Implementation

In practice, the implementation of the schemes written as in (3.6) and (3.10) leads to two numerical issues. First, the matrix that should be inverted is ill-conditioned for small ε. This leads to large numerical errors near the diffusive regime and deteriorates the AP property of the schemes. Second, the evaluation of integrated differences that are of the same order, such as $\|f^n - \mu_f\mathcal{M}\|_{2,\mathcal{M}}$ when n is large leads to the accumulations of machine epsilons and poor approximations of these quantities.

In this section, we present an equivalent formulation of the schemes (3.6) and (3.10) which solves both these issues. It is based on a perturbative micro-macro formulation introduced in

Section 4.1. In the next subsection, we justify the uniform conditioning of the reformulated linear systems on the whole range of ε (including $\varepsilon = 0$). In Section 4.2, we write the linear systems in explicit matrix form so that the reader may implement the scheme easily and replicate the numerical results of Section 5. The test cases illustrate the theoretical results of Theorem 1 and Theorem 2, namely the asymptotic accuracy of our schemes in both the diffusion limit and the large time asymptotics.

4.1. **Micro-macro formulations of the schemes.** Given $(f_{ij}^n)_{i\in\mathcal{I},j\in\mathcal{J},n\in\mathbb{N}}$, let us introduce the micro and macro unknowns, $(h_{ij}^n)_{i\in\mathcal{I},j\in\mathcal{J},n\in\mathbb{N}}$ and $(\lambda_i^n)_{i\in\mathcal{I},n\in\mathbb{N}}$ respectively, which satisfy for all $i \in \mathcal{I}$, $j \in \mathcal{J}$ and $n \in \mathbb{N}$

$$f_{ij}^n = \mu_f \, \mathcal{M}_j + \lambda_i^n \, \mathcal{M}_j + \varepsilon \, h_{ij}^n \, \mathcal{M}_j \,, \tag{4.1}$$

$$\lambda_i^n = \rho_i^n - \mu_f = \sum_{j\in\mathcal{I}} f_{ij}^n \, \Delta v_j - \mu_f \,. \tag{4.2}$$

An immediate consequence of these expressions is that

$$\sum_{j\in\mathcal{J}} h_{ij}^n \, \mathcal{M}_j \, \Delta v_j \;=\; 0 \qquad \forall i \in \mathcal{I} \,, \forall n \in \mathbb{N} \,. \tag{4.3}$$

If one writes the counterpart of the continuity equation (3.21) with the new unknowns, one gets the following macroscopic evolution equation

$$\Delta x_i (\lambda_i^{n+1} - \lambda_i^n) \,+\, \Delta t \sum_{j\in\mathcal{J}} \Delta v_j \, v_j \, \mathcal{M}_j \, \frac{h_{i+1,j}^{n+1} - h_{i-1,j}^{n+1}}{2} \;=\; 0 \,, \qquad \forall i \in \mathcal{I} \,. \tag{4.4}$$

- *In the Fokker-Planck case*, if one plugs (4.1) into the scheme (3.6) and then subtracts (4.4) multiplied by $-\Delta v_j \, \varepsilon$, one obtains the following microscopic evolution equation

$$\begin{aligned} &\varepsilon^2 \, \Delta x_i \, \Delta v_j \, (h_{ij}^{n+1} - h_{ij}^n) \,+\, \Delta t \, \Delta v_j \, v_j \frac{\lambda_{i+1}^{n+1} - \lambda_{i-1}^{n+1}}{2} \\ &\quad + \varepsilon \, \Delta t \, \Delta v_j \left(v_j \, \frac{h_{i+1,j}^{n+1} - h_{i-1,j}^{n+1}}{2} - \sum_{k\in\mathcal{J}} \Delta v_k \, v_k \, \mathcal{M}_k \, \frac{h_{i+1,k}^{n+1} - h_{i-1,k}^{n+1}}{2} \right) \\ &\quad = \; \Delta t \, \Delta x_i \left(\frac{\mathcal{M}_{j+1/2}^*}{\Delta v_{j+1/2} \, \mathcal{M}_j} \left(h_{i,j+1}^{n+1} - h_{ij}^{n+1} \right) - \frac{\mathcal{M}_{j-1/2}^*}{\Delta v_{j-1/2} \, \mathcal{M}_j} \left(h_{ij}^{n+1} - h_{i,j-1}^{n+1} \right) \right) . \end{aligned} \tag{4.5}$$

- *In the BGK case*, if one plugs (4.1) into the scheme (3.10), one obtains the following microscopic evolution equation

$$\begin{aligned} &\varepsilon^2 \, \Delta x_i \, \Delta v_j \, (h_{ij}^{n+1} - h_{ij}^n) \,+\, \Delta t \, \Delta v_j \, v_j \frac{\lambda_{i+1}^{n+1} - \lambda_{i-1}^{n+1}}{2} \\ &\quad + \varepsilon \, \Delta t \, \Delta v_j \left(v_j \, \frac{h_{i+1,j}^{n+1} - h_{i-1,j}^{n+1}}{2} - \sum_{k\in\mathcal{J}} \Delta v_k \, v_k \, \mathcal{M}_k \, \frac{h_{i+1,k}^{n+1} - h_{i-1,k}^{n+1}}{2} \right) \\ &\quad = \; -\Delta t \, \Delta x_i \, \Delta v_j h_{ij}^{n+1} \,. \end{aligned} \tag{4.6}$$

Based on the previous computations, one can define three schemes. In all three cases we start from the given initial data $(f_{ij}^0)_{i,j} \in \mathbb{R}^{\mathcal{I}\times\mathcal{J}}$ and define $(h_{ij}^0)_{i,j} \in \mathbb{R}^{\mathcal{I}\times\mathcal{J}}$ and $(\lambda_i^0)_i \in \mathbb{R}^{\mathcal{I}}$ according to (4.1) and (4.2). If $\varepsilon = 0$, we define $h_{ij}^0 = 0$ for all i, j. Then the schemes are as follows.

($\mathbf{S}_f^\varepsilon$) The unknown $(f_{ij}^n)_{i,j,n} \in \mathbb{R}^{\mathcal{I}\times\mathcal{J}\times\mathbb{N}}$ solves (3.6) in the Fokker-Planck case and (3.10) in the BGK case. The micro and macro unknowns $(h_{ij}^n)_{i,j,n} \in \mathbb{R}^{\mathcal{I}\times\mathcal{J}\times\mathbb{N}}$ and $(\lambda_i^n)_{i,n} \in \mathbb{R}^{\mathcal{I}\times\mathbb{N}}$ are defined by (4.1) and (4.2).

($\mathbf{S}^\varepsilon_{\lambda,h}$) The micro and macro unknowns $(h^n_{ij})_{i,j,n}$ and $(\lambda^n_i)_{i,n}$ solve (4.4), (4.5) in the Fokker-Planck case and (4.4), (4.6) in the BGK case. The unknown $(f^n_{ij})_{i,j,n}$ is defined by (4.1).

($\tilde{\mathbf{S}}^\varepsilon_{\lambda,h}$) The micro and macro unknowns $(h^n_{ij})_{i,j,n}$ and $(\lambda^n_i)_{i,n}$ solve the overdetermined linear system (4.3), (4.4), (4.5) in the Fokker-Planck case and (4.3), (4.4), (4.6) in the BGK case. The unknown $(f^n_{ij})_{i,j,n}$ is defined by (4.1).

These schemes happen to be well-posed and equivalent in the kinetic regime $\varepsilon > 0$ but not in the diffusive regime $\varepsilon = 0$.

Proposition 4. *The three versions of the schemes satisfy the following properties.*

(i) *If $\varepsilon > 0$, $(\mathbf{S}^\varepsilon_f)$, $(\mathbf{S}^\varepsilon_{\lambda,h})$ and $(\tilde{\mathbf{S}}^\varepsilon_{\lambda,h})$ define the same unique $(f^n_{ij})_{i,j,n}$, $(h^n_{ij})_{i,j,n}$ and $(\lambda^n_i)_{i,n}$.*

(ii) *If $\varepsilon = 0$, the equivalence between the schemes does not hold. More precisely, $(\mathbf{S}^0_f)$ is ill-posed, $(\mathbf{S}^0_{\lambda,h})$ is well-posed in the BGK case and ill-posed in the Fokker-Planck case and $(\tilde{\mathbf{S}}^0_{\lambda,h})$ is well-posed.*

Proof.

(i) The computations of the beginning of this section already showed that $(\mathbf{S}^\varepsilon_f) \Rightarrow (\mathbf{S}^\varepsilon_{\lambda,h})$. Let us now prove that $(\mathbf{S}^\varepsilon_{\lambda,h}) \Rightarrow (\tilde{\mathbf{S}}^\varepsilon_{\lambda,h})$. At $n = 0$ one has $\sum_{j\in\mathcal{J}} \Delta v_j\, h^0_{ij}\, \mathcal{M}_j \;=\; 0$ for all $i \in \mathcal{I}$. In the Fokker-Planck case, one can multiply (4.5) by $\mathcal{M}_j$ and sum over $j \in \mathcal{J}$ to obtain

$$\sum_{j\in\mathcal{J}} \Delta v_j\, (\varepsilon^2\, h^{n+1}_{ij} - \varepsilon^2\, h^n_{ij})\, \mathcal{M}_j \;=\; 0\,, \qquad \forall i \in \mathcal{I}\,, \quad \forall n \in \mathbb{N}\,.$$

In the BGK case, it yields

$$\sum_{j\in\mathcal{J}} \Delta v_j\, ((\varepsilon^2 + \Delta t)\, h^{n+1}_{ij} - \varepsilon^2\, h^n_{ij})\, \mathcal{M}_j \;=\; 0\,, \qquad \forall i \in \mathcal{I}\,, \quad \forall n \in \mathbb{N}\,. \tag{4.7}$$

Therefore (4.3) holds. Finally let us show that $(\tilde{\mathbf{S}}^\varepsilon_{\lambda,h}) \Rightarrow (\mathbf{S}^\varepsilon_f)$. Since (4.3) holds, by multiplying (4.1) by Δv_j and summing over $j \in \mathcal{J}$ one obtains (4.2). Then, just add (4.4) multiplied by $\varepsilon\Delta v_j$ to (4.5) (*resp.* (4.6)) to obtain (3.6) (*resp.* (3.10)).

(ii) Let $\varepsilon = 0$. In the case of the Fokker-Planck operator, let us define the discrete collision matrix Q^δ_{FP} satisfying for all $g = (g_j)_{j\in\mathcal{J}}$

$$Q^\delta_{FP}\, g \;:=\; \left(\frac{1}{\Delta v_j} \left(\frac{\mathcal{M}^*_{j+1/2}}{\Delta v_{j+1/2}\, \mathcal{M}_j}\, (g_{j+1} - g_j) - \frac{\mathcal{M}^*_{j-1/2}}{\Delta v_{j-1/2}\, \mathcal{M}_j}\, (g_j - g_{j-1}) \right) \right)_{j\in\mathcal{J}}\,.$$

Since one has $\left\langle Q^\delta_{FP}\, g,\, g \right\rangle_{2,\mathcal{M}} = \|D_v g\|^2_{2,\mathcal{M}}$, it is clear that $\mathrm{Ker}(Q^\delta_{FP}) = \mathrm{Span}\{(1,1,\dots,1)\}$ and therefore the linear system (4.5) (with $\varepsilon = 0$) has infinitely many solutions since $(v_j)_{j\in\mathcal{J}} \in \mathrm{Ker}(Q^\delta_{FP})^\perp = \{g,\ \text{s. t.}\ \sum_{j\in\mathcal{J}} g_j \mathcal{M}_j \Delta v_j \;=\; 0\}$. If one multiplies (4.5) by $v_j \mathcal{M}_j$ and sums over j, one gets for any solution

$$\sum_{j\in\mathcal{J}} \Delta v_j\, v_j\, \mathcal{M}_j\, h^{n+1}_{ij} \;=\; -m_2^{\Delta v}\, \frac{\lambda^{n+1}_{i+1} - \lambda^{n+1}_{i-1}}{2\Delta x_i}\,.$$

Then by plugging this into (4.4), we infer that $(\lambda^{n+1}_i)_{i\in\mathcal{I}}$ is determined uniquely since it satisfies our limit heat equation (3.23). Hence, the linear system of $(\mathbf{S}^0_{\lambda,h})$ has infinitely many solutions. If we now consider the scheme $(\tilde{\mathbf{S}}^0_{\lambda,h})$ one has the additional equations $\sum_{j\in\mathcal{J}} h_{ij}\, \mathcal{M}_j\, \Delta v_j \;=\; 0$ for all $i \in \mathcal{I}$ which correspond to $(h_{ij})_j \in \mathrm{Ker}(Q^\delta_{FP})^\perp$ for all $i \in \mathcal{I}$. The solution is now unique which proves well-posedness of $(\tilde{\mathbf{S}}^0_{\lambda,h})$.

In the case of the BGK operator, the situation is simpler since the discrete collision operator Q^δ_{BGK} is a diagonal positive definite matrix. Hence both $(\mathbf{S}^0_{\lambda,h})$ and $(\tilde{\mathbf{S}}^0_{\lambda,h})$ are well-posed. With similar arguments than those developed in the foregoing proof, it is easy to show that $(\mathbf{S}^0_f)$ is ill-posed in both the Fokker-Planck and BGK cases.

□

An immediate consequence of the previous proposition is the following

Corollary 2. *The condition number of the linear system in $(\tilde{\mathbf{S}}^\varepsilon_{\lambda,h})$ is uniformly bounded with respect to $\varepsilon \in [0,1]$.*

Proof. By invertibility of the matrices, the condition number is well defined on the whole range $\varepsilon \in [0,1]$. Moreover it is clearly continuous in ε. Hence the boundedness follows from uniform continuity. □

Remark 7. *By taking our schemes for the kinetic equations at $\varepsilon = 0$ in their formulations $(\tilde{\mathbf{S}}^0_{\lambda,h})$, one is exactly solving the discrete heat equation (3.23) for $(\lambda_i)_{i\in\mathcal{I}}$, as expected in the diffusion limit.*

4.2. **Explicit matrix formulation.** Following the results of the previous section, our schemes should be implemented in their robust form $(\tilde{\mathbf{S}}^\varepsilon_{\lambda,h})$. In this section, we write the explicit matrix formulation of this scheme. At time t^n, both schemes rewrite in linear system form

$$\mathbb{M}^\varepsilon_{\rm col}\,\mathbf{U}^{n+1} \;=\; \mathbb{D}^\varepsilon\,\mathbf{U}^n$$

with col $=$ FP or BGK. The vector $\mathbf{U}^n$ has two sets of components corresponding to the macro and micro variable. They are indexed by control volumes. The first set is $\mathbf{U}^n_{\mathcal{X}_i} = \lambda^n_i$ for $i \in \mathcal{I}$ and the second is $\mathbf{U}^n_{K_{ij}} = h^n_{ij}$ with $(i,j) \in \mathcal{I}\times\mathcal{J}$. From the knowledge of $\mathbf{U}^n$, one can compute the usual distribution function $(f^n_{ij})_{i,j}$ thanks to (4.1). Since the linear system is rectangular (but has a unique solution as shown previously), it can be solved thanks to the pseudo-inverse

$$\mathbf{U}^{n+1} \;=\; \left[(\mathbb{M}^\varepsilon_{\rm col})^\top\,\mathbb{M}^\varepsilon_{\rm col}\right]^{-1}(\mathbb{M}^\varepsilon_{\rm col})^\top\,\mathbb{D}^\varepsilon\,\mathbf{U}^n\,. \tag{4.8}$$

This linear system is a rewriting of the relations (4.3), (4.4), (4.5) in the Fokker-Planck case (or (4.6) in the BGK case). The corresponding matrices $\mathbb{M}^\varepsilon_{\rm col}$ are defined by

$$\mathbb{M}^\varepsilon_{\rm col} \;=\; \begin{pmatrix} \mathbb{I}_N & \mathbb{M}^{\lambda,h} \\ \mathbb{M}^{h,\lambda} & \mathbb{M}^{h,h}_{\rm col} \\ 0 & \mathbb{N}^{h,h} \end{pmatrix}, \qquad \mathbb{D}^\varepsilon \;=\; \begin{pmatrix} \mathbb{I}_N & 0 \\ 0 & \varepsilon^2\mathbb{I}_{2L} \\ 0 & 0 \end{pmatrix}.$$

Here, we denoted by $\mathbb{I}_K$ the $K\times K$ identity matrix and the other matrices are defined as follows. The matrix $\mathbb{M}^{\lambda,h}$ has non-zero coefficients

$$\mathbb{M}^{\lambda,h}_{\mathcal{X}_i,K_{i+1,j}} \;=\; \frac{v_j\,\mathcal{M}_j\,\Delta v_j\,\Delta t}{2\,\Delta x_i}\,, \qquad \mathbb{M}^{\lambda,h}_{\mathcal{X}_i,K_{i-1,j}} \;=\; -\frac{v_j\,\mathcal{M}_j\,\Delta v_j\,\Delta t}{2\,\Delta x_i}\,, \quad \forall (i,j)\in\mathcal{I}\times\mathcal{J}\,.$$

The matrix $\mathbb{M}^{h,\lambda}$ has non-zero coefficients

$$\mathbb{M}^{h,\lambda}_{K_{i,j},\mathcal{X}_{i+1}} \;=\; \frac{v_j\,\Delta t}{2\,\Delta x_i}\,, \qquad \mathbb{M}^{h,\lambda}_{K_{i,j},\mathcal{X}_{i-1}} \;=\; -\frac{v_j\,\Delta t}{2\,\Delta x_i}\,, \quad \forall (i,j)\in\mathcal{I}\times\mathcal{J}\,.$$

The matrix $\mathbb{M}^{h,h}_{BGK}$ has non-zero coefficients

$$\forall (i,j) \in \mathcal{I} \times \mathcal{J} \quad \forall k \in \mathcal{J}\,,\ k \neq j\,, \qquad \mathbb{M}^{h,h}_{K_{ij},K_{ij}} \ = \ \varepsilon^2 + \Delta t\,,$$

$$\mathbb{M}^{h,h}_{K_{ij},K_{i+1,j}} \ = \ \varepsilon\,\frac{v_j\,\Delta t\,(1-\Delta v_j\,\mathcal{M}_j)}{2\,\Delta x_i}\,, \qquad \mathbb{M}^{h,h}_{K_{ij},K_{i-1,j}} \ = \ -\,\varepsilon\,\frac{v_j\,\Delta t\,(1-\Delta v_j\,\mathcal{M}_j)}{2\,\Delta x_i}\,,$$

$$\mathbb{M}^{h,h}_{K_{ij},K_{i+1,k}} \ = \ -\,\varepsilon\,\frac{v_k\,\Delta t\,\Delta v_k\,\mathcal{M}_k}{2\,\Delta x_i}\,, \qquad \mathbb{M}^{h,h}_{K_{ij},K_{i-1,k}} \ = \ \varepsilon\,\frac{v_k\,\Delta t\,\Delta v_k\,\mathcal{M}_k}{2\,\Delta x_i}\,.$$

The matrix $\mathbb{M}^{h,h}_{FP}$ has the same coefficients as $\mathbb{M}^{h,h}_{BGK}$ except for the following (outside the boundary in velocity)

$$\forall (i,j) \in \mathcal{I} \times \mathcal{J}^*\,, \qquad \mathbb{M}^{h,h}_{K_{ij},K_{ij}} \ = \ \varepsilon^2 + \frac{\Delta t}{\Delta v_j \mathcal{M}_j}\left(\frac{\mathcal{M}^*_{j+1/2}}{\Delta v_{j+1/2}} + \frac{\mathcal{M}^*_{j-1/2}}{\Delta v_{j-1/2}}\right),$$

$$\mathbb{M}^{h,h}_{K_{ij},K_{i,j+1}} \ = \ -\,\frac{\Delta t\,\mathcal{M}^*_{j+1/2}}{\Delta v_j\,\mathcal{M}_j\,\Delta v_{j+1/2}}\,, \qquad \mathbb{M}^{h,h}_{K_{ij},K_{i,j-1}} \ = \ -\,\frac{\Delta t\,\mathcal{M}^*_{j-1/2}}{\Delta v_j\,\mathcal{M}_j\,\Delta v_{j-1/2}}\,.$$

At the boundary, $\forall i \in \mathcal{I}$,

$$\mathbb{M}^{h,h}_{K_{i,-L+1},K_{i,-L+1}} \ = \ \varepsilon^2 + \frac{\Delta t}{\Delta v_{-L+1}\mathcal{M}_{-L+1}}\left(\frac{\mathcal{M}^*_{-L+3/2}}{\Delta v_{-L+3/2}}\right),$$

$$\mathbb{M}^{h,h}_{K_{i,-L+1},K_{i,-L+2}} \ = \ -\,\frac{\Delta t\,\mathcal{M}^*_{-L+3/2}}{\Delta v_{-L+1}\,\mathcal{M}_{-L+1}\,\Delta v_{-L+3/2}}\,,$$

$$\mathbb{M}^{h,h}_{K_{i,L},K_{i,L}} \ = \ \varepsilon^2 + \frac{\Delta t}{\Delta v_L \mathcal{M}_L}\left(\frac{\mathcal{M}^*_{L-1/2}}{\Delta v_{L-1/2}}\right), \qquad \mathbb{M}^{h,h}_{K_{i,L},K_{i,L-1}} \ = \ -\,\frac{\Delta t\,\mathcal{M}^*_{L-1/2}}{\Delta v_L\,\mathcal{M}_L\,\Delta v_{L-1/2}}\,.$$

Finally the matrix $\mathbb{N}^{h,h}$, which comes from (4.3), has non-zero coefficients

$$\mathbb{N}^{h,h}_{\mathcal{X}_i,K_{i,j}} \ = \ \mathcal{M}_j\,\Delta v_j, \qquad \forall (i,j) \in \mathcal{I} \times \mathcal{J}\,.$$

5. Numerical simulations

In this final section, we shall present simulations for the overdetermined micro-macro versions $(\tilde{\mathbf{S}}^{\varepsilon}_{\lambda,h})$ of our numerical schemes. We will address numerically all the features of our numerical methods, namely their AP properties in a first subsection, and the discrete hypocoercivity in the next ones, along with some fine properties of the numerical solutions.

The numerical method has been written in *Python 3.6*, using *Jupyter Lab*, and is available, together with all the test cases presented in this section in [6].

Unless stated otherwise, in all the simulations we will take $L = 20$, namely 41 points in a velocity box $[-v_*, v_*]$, with $v_* = 8$, and $N = 51$ points in the torus $[0,1]$. Given that our schemes are unconditionally stable, we shall choose $\Delta t = 0.1$.

5.1. **Numerical investigation of the AP Property.**

Test 1. Convergence toward the heat equation. Let us start by investigating numerically the AP property of our schemes. We choose as an initial condition the following smooth, far-from-equilibrium distribution

$$f_0(x,v) := \frac{1}{\sqrt{2\pi}}\,v^4\,e^{-v^2/2}\,\frac{1+\cos(4\pi x)}{2}, \quad \forall x \in \mathbb{T}, v \in [-v_*, v_*]. \tag{5.1}$$

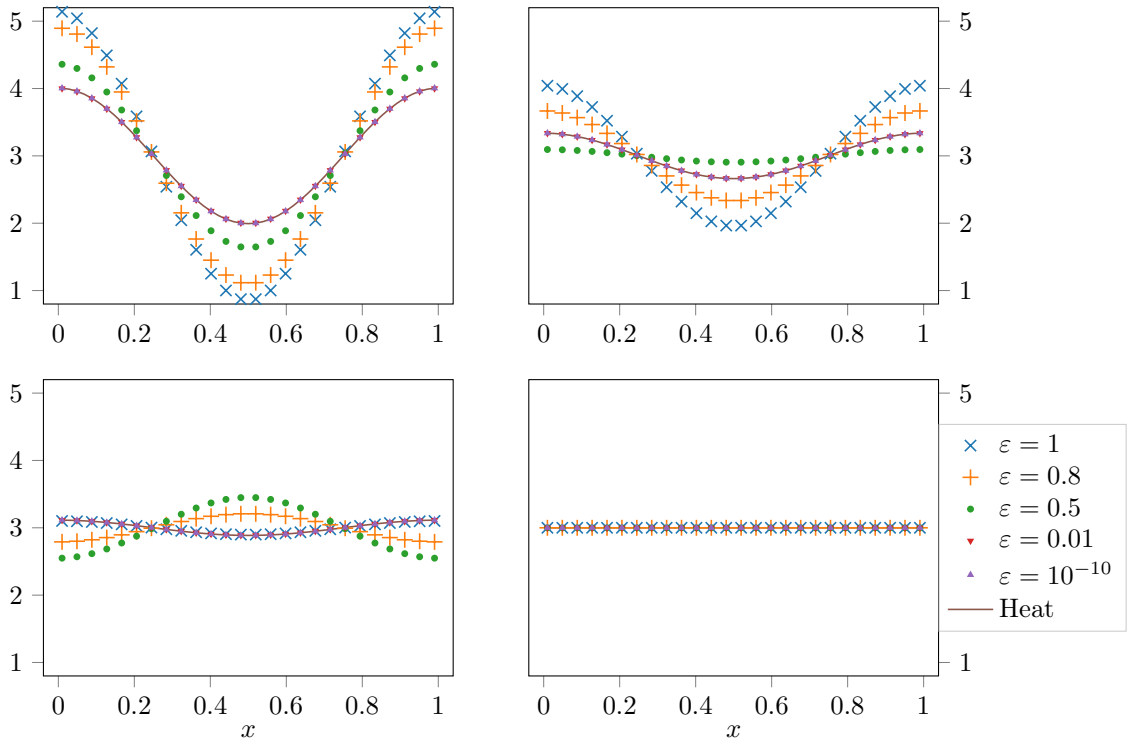


FIGURE 3. **Test 1.** Comparison of the solution to the heat equation obtained from the scheme (3.23) (solid line) with the solution obtained with the **Fokker-Planck** scheme (4.5) with different ε, at times $t = 0.05$, 0.1, 0.15 and 10.

We solve first the kinetic Fokker-Planck equation with scheme (4.5), for different values of ε, and compare the discrete density ρ^ε at different times with the solution to the discrete heat equation given by (3.23). The results can be found in Figure 3, where we observe a good agreement between the convergence result of Theorem 1 and our simulations.

We then perform the same numerical investigations for the linear BGK equation with scheme (4.6). Again, we observe in Figure 4 the convergence with respect to ε of the density ρ^ε towards the solution to the discrete heat equation ρ, as predicted by Theorem 1.

We will now investigate the hypocoercivity properties of our schemes, with various numerical experiments in the Fokker-Planck and BGK cases.

5.2. **The Fokker-Planck case.**

Test 2. Trend to equilibrium. We shall fix $\varepsilon = 1$ for this numerical test case.

Let us first start to verify the results of Theorem 2, namely that the discrete solution f^ε to the kinetic Fokker-Planck equation converges towards $\mu_f \mathcal{M}$ in the discrete, weighted, L^2 topology, at an exponential rate.

We first perform this experiment with an initial data chosen to be random, uniformly distributed on the $(x, v)-$plane, but supported in $v \in [-3, 3]$ (also shown in the left side of Figure 5). We observe in the rightmost part of Figure 5 that the numerical rate of convergence of the discrete norm $\|f - \mu_f \mathcal{M}\|_{2,\gamma}$ obtained with the scheme (4.5) is indeed exponential. It also matches the rate of convergence in the spatial L^2 norm of the density ρ^ε towards the global

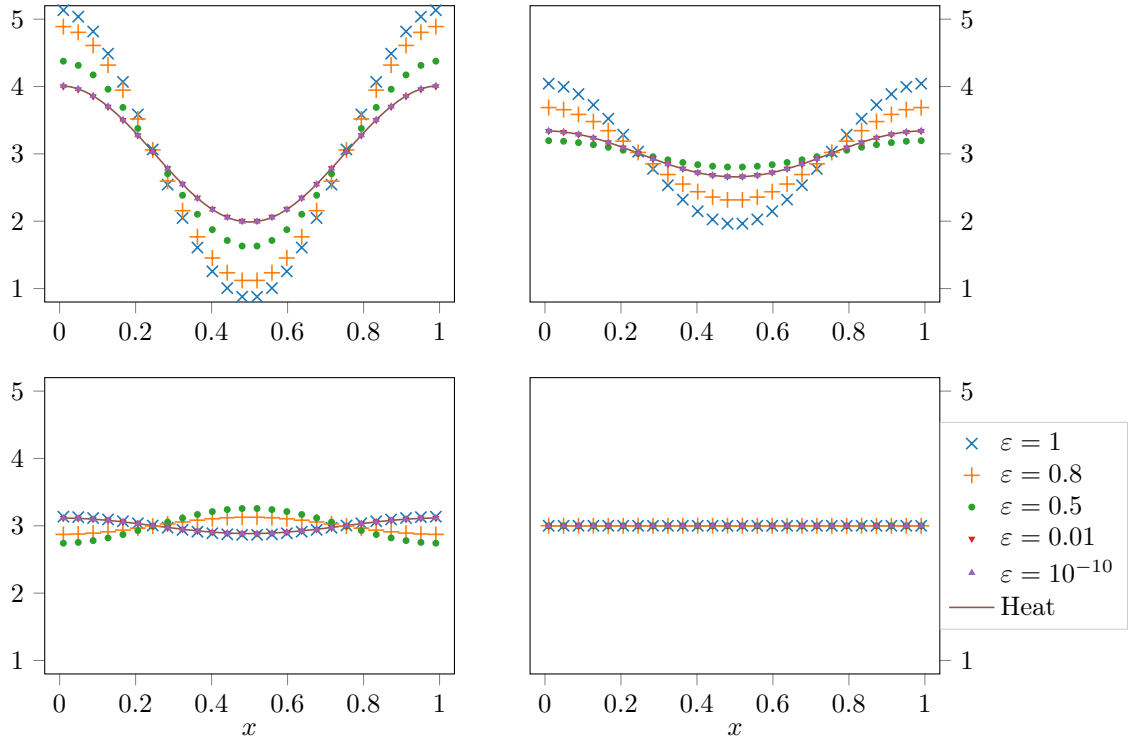


FIGURE 4. **Test 1.** Comparison of the solution to the heat equation obtained from the scheme (3.23) (solid line) with the solution obtained with the **BGK** scheme (4.6) with different ε, at times $t = 0.05$, 0.1, 0.15 and 10.

mass μ_f. The numerical hypocoercivity rate for this test case is 1.003, which is close to the rate observed[1] in the paper [22].

These observations are confirmed by choosing another initial condition, such as the indicator of a ball in the $(x, v)-$phase plane, as seen on the left of Figure 6. We observe in the rightmost part of that figure the same behavior than in the previous case. The numerical rate of decay is now 1.96, quite similar to the value observed in [22].

Finally, we also present in Figure 7 a snapshot of the particle distribution function in the $(x, v)-$phase-plane at different times, illustrating its convergence towards the global Maxwellian distribution.

Test 3. Dependency of the exponential rate on ε. For this final Fokker-Planck test, we study the influence of ε on the rate of exponential decay of the quantity $\|f^\varepsilon - \mu_f \mathcal{M}\|_{2,\gamma}$. We choose initially a uniformly distributed distribution f_0 in the $(x, v)-$plane, with values between 0 and 1, and we let ε vary from the purely kinetic regime $\varepsilon = 1$ to the purely diffusive regime $\varepsilon = 0$, without ever changing the other grid parameters.

We represent the results of this numerical investigation in Figure 8, by plotting the different numerical values of $\|f^\varepsilon(t) - \mu_f \mathcal{M}\|_{2,\gamma}$. We observe that the rate of exponential decay seems to have a lower bound which is uniform on ε, illustrating the fact that the result of Theorem 2 is

[1]Note that the quantity we compute in our paper is the square root of the one computed in [22], explaining the factor 2 between the observations.

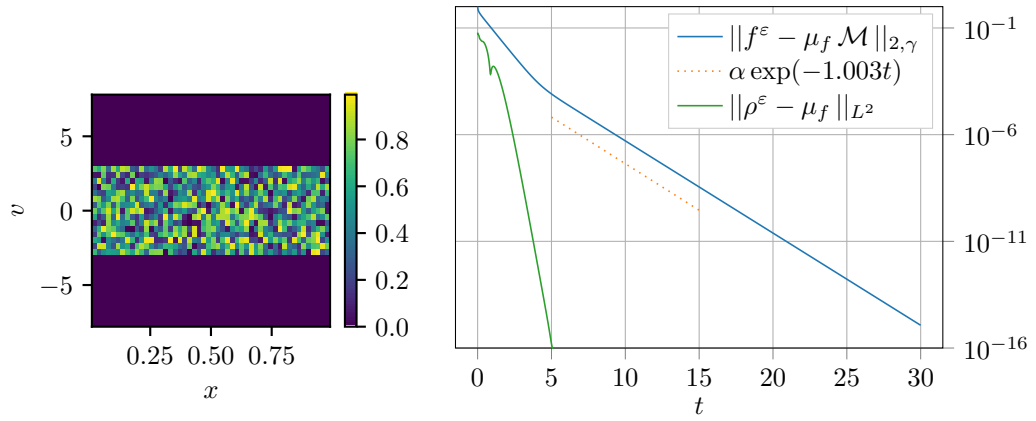


FIGURE 5. **Test 2.** Left: Initial data in the $(x, v)-$phase-plane. Right: Time evolution of the weighted L^2 norm of the difference between f^ε and the global equilibrium, $\varepsilon = 1$ in the **Fokker-Planck** case, and observed exponential rate.

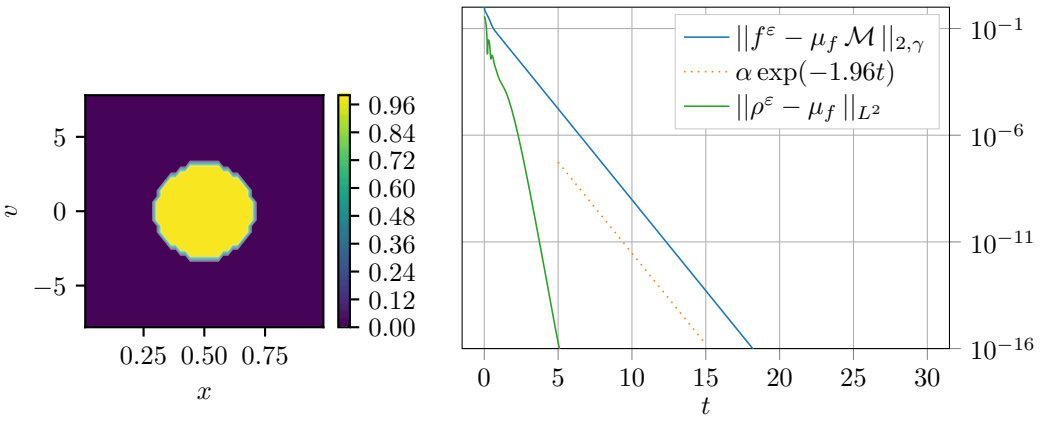


FIGURE 6. **Test 2.** Left: Initial data in the $(x, v)-$phase-plane. Right: Time evolution of the weighted L^2 norm of the difference between f^ε and the global equilibrium, $\varepsilon = 1$, **Fokker-Planck** case.

suboptimal concerning the rate β obtained. This could of course be improved by tracking exactly the dependence in ε in the estimates used in the proof of Theorem 2.

To make this latter statement more explicit, we show in Table 1 the value of the rate of decay, as a function of ε. This rate decreases with ε towards a limit value.

5.3. The BGK case.

Test 4. Oscillations during the macroscopic relaxation to equilibrium. We now consider the question of the macroscopic relaxation towards equilibrium. So far, we investigated the behavior of the rate of relaxation in $L^2(\mathrm{d}x\mathrm{d}\gamma)$ of f^ε towards its global equilibrium $\mu_f\mathcal{M}$, namely kinetic relaxation. We will now turn to a numerical study of the rate of relaxation in $L^2(\mathrm{d}x)$ of the velocity averages of f^ε, namely ρ^ε, towards their constant-in-space global equilibrium μ_f. We shall call this later case *macroscopic* relaxation.

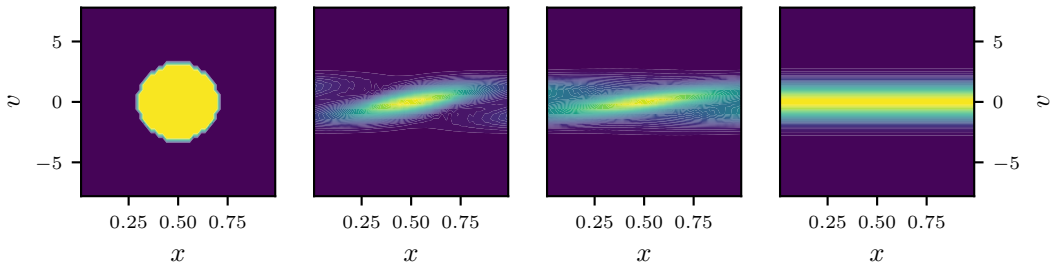


FIGURE 7. **Test 2.** Snapshot of the particle distribution function $f^\varepsilon(t,x,v)$ in the $(x,v)-$phase-plane, at times $t = 0$, 0.3, 0.6, and 30.

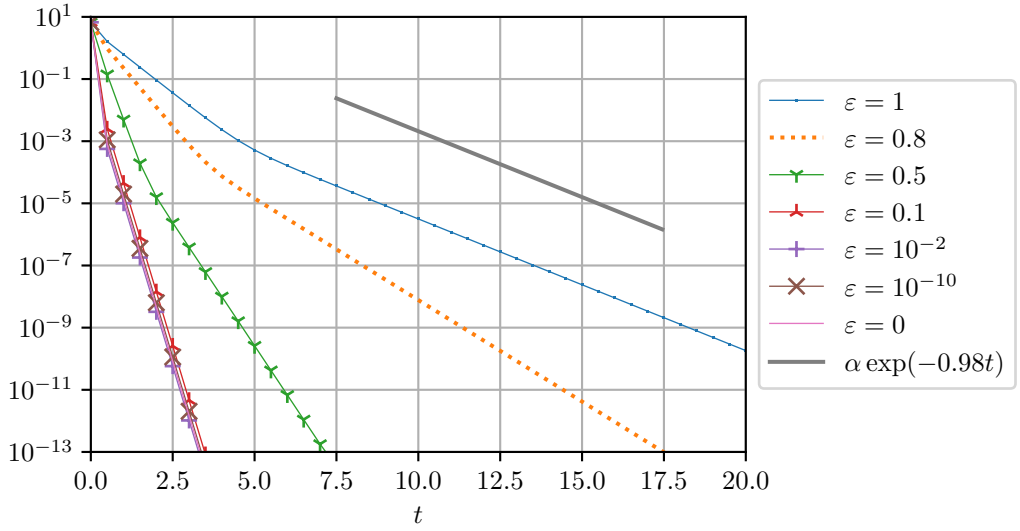


FIGURE 8. **Test 3.** Comparison of the rate of convergence of $\|f^\varepsilon - \mu_f \mathcal{M}\|_{2,\gamma}$ for different values of ε, in the **Fokker-Planck** case.

ε	Rate
1	-0.98
0.8	-1.5
0.5	-3.65
10^{-1}	-8.04
10^{-2}	-8.05
10^{-10}	-8.05
0	-8.05

TABLE 1. **Test 3.** Rate of exponential decay of $\|f^\varepsilon(t) - \mu_f \mathcal{M}\|_{2,\gamma}$ with respect to ε.

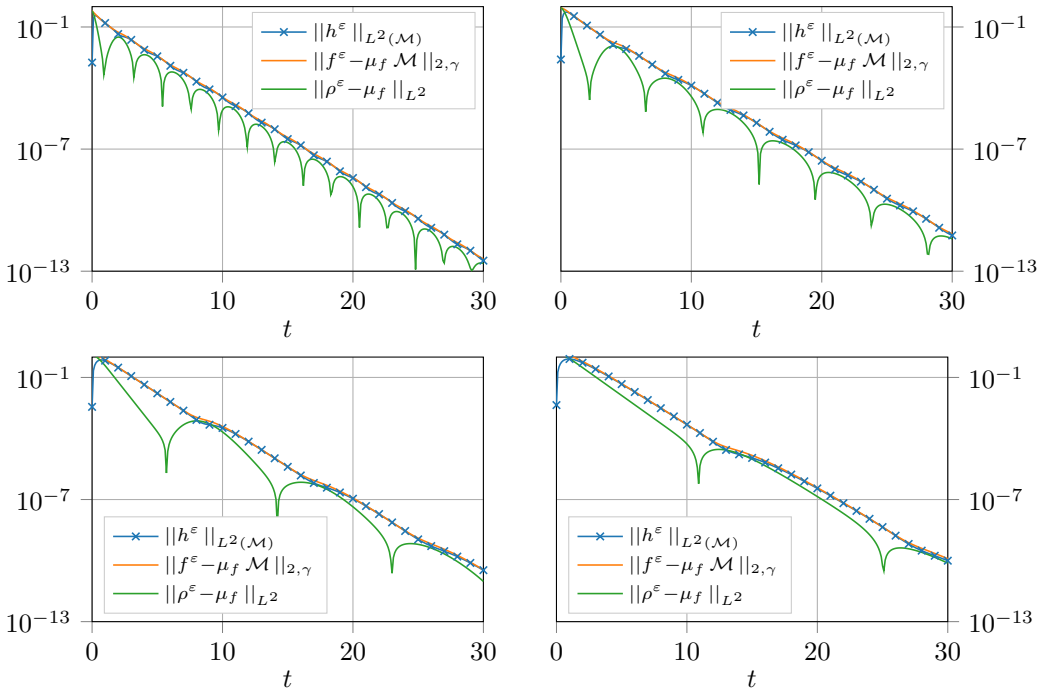


FIGURE 9. **Test 4.** Influence of the torus size R on the macroscopic relaxation, for $R = \pi/4$, $\pi/2$, π, and $3\pi/2$, in the **BGK** case.

The behavior of this macroscopic quantity has already been presented in Figures 5 and 6, where one can notice that the macroscopic relaxation seems bounded by the kinetic one. Nevertheless, this macroscopic behavior is actually richer, and can even be not monotonous. To study it more precisely, we shall consider the following close-to-equilibrium initial condition

$$f_0(x,v) := \left(1 + \cos\left(\frac{2\pi x}{R}\right)\right)\mathcal{M}(v), \quad \forall x \in [0,R], \quad v \in [-v_*, v_*]. \tag{5.2}$$

Our goal is to change the size R of the torus, to see how it affects the behavior of the macroscopic relaxation. Note that in the previous test cases, R was equal to 1.

Figure 9 presents, for the BGK case, the time evolution of the local relaxation $\|f^\varepsilon(t) - \mu_f\mathcal{M}\|_{2,\gamma}$, and of the macroscopic relaxation $\|\rho^\varepsilon(t) - \mu_f\|_{L^2(\mathrm{d}x)}$, for 4 box sizes R. The $L^2(\mathcal{M}\mathrm{d}x\mathrm{d}v)$ norm of the scaled, microscopic unknown $h^\varepsilon := (f^\varepsilon - \mu_f\mathcal{M})/\mathcal{M}$ is also shown on these figures. The scaling constant ε is chosen equal to 1 in all these simulations.

We observe that, as expected through Theorem 2, the local relaxation rate is monotone and exponential. This rate is inversely proportional to the box size R. Nevertheless, the macroscopic rate of relaxation is not monotone, exhibiting exponentially damped oscillations. We also observe that the upper envelope of these oscillations seems to be given by $\|h^\varepsilon(t)\|_{L^2(\mathcal{M}\mathrm{d}x\mathrm{d}v)}$, and that this quantity seems to decay in a monotone but slightly oscillating way.

Inspired by the analysis conducted in [25], we will give a simple interpretation for these oscillatory behaviors. Since the BGK equation we consider is linear, the initial value (5.2) we chose corresponds to observing the time evolution of the Fourier mode in x of frequency $k = 2\pi/R$.

Torus length R	Oscillation period ν^{-1}	Speed $R\nu$
$\pi/4$	2.31	0.340
$\pi/2$	4.33	0.363
π	8.67	0.362
$3\pi/2$	13.5	0.350

TABLE 2. **Test 4.** Influence of the torus length on the oscillation period of $\|\rho^\varepsilon - \mu_f\|_{L^2}$, in the **BGK** case.

Going to Fourier in x in the BGK equation (1.1) amounts to solve the following integral equation

$$\partial_t g = \left(\int_{\mathbb{R}} g(v)\,\mathrm{d}v\right)\mathcal{M}(v) - (1+ikv)\,g. \tag{5.3}$$

The right hand side of (5.3) is the sum of an operator of multiplication by $\mu(v) := -(1+ikv)$ and of a rank-1 projector (towards the global Maxwellian), that we shall denote by $\mathcal{K}$. This operator then generates a strongly continuous semigroup $T_k(t)$. Moreover, according to Weyl's Theorem [38], its spectrum in L^2 contains the spectrum of the multiplication operator, namely the numerical range of μ:

$$\{-1+iy : y\in\mathbb{R}\},$$

together with some discrete eigenvalues located in the complex plane on the right of this set. Since we proved in Theorem 2 that the linear BGK equation is hypocoercive with a rate at least $\beta/2$, necessarily, such a discrete nonzero eigenvalue $\lambda\in\mathbb{C}^*$ verify

$$-1 \le \Re e\,(\lambda) \le -\beta/2.$$

Let $g_\lambda \in L^2(\mathrm{d}x\mathrm{d}\gamma)\setminus\{0\}$ be the associated eigenvector. One then has

$$\mathcal{K} g_\lambda = (1+\lambda+ikv)\,g_\lambda. \tag{5.4}$$

Taking the $L^2(\mathrm{d}x\mathrm{d}\gamma)-$inner product of this relation with g_λ yields

$$\begin{aligned}\lambda &= -1 + \frac{\langle \mathcal{K} g_\lambda, g_\lambda\rangle_{L^2(\mathrm{d}x\mathrm{d}\gamma)}}{\|g_\lambda\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}} - ik\frac{\int_{\mathbb{R}} v\,|g_\lambda|^2\,\mathrm{d}\gamma}{\|g_\lambda\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}}\\ &= -\left[1 - \frac{\left(\int_{\mathbb{R}} g_\lambda(v)\,\mathrm{d}v\right)^2}{\|g_\lambda\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}}\right] - ik\frac{\int_{\mathbb{R}} v\,|g_\lambda|^2\,\mathrm{d}\gamma}{\|g_\lambda\|^2_{L^2(\mathrm{d}x\mathrm{d}\gamma)}}.\end{aligned} \tag{5.5}$$

Note that the case of the space-homogeneous solution to (1.1) corresponds to $k=0$, so that taking the integral of (5.4) yields that g_λ is of zero mass. Hence, in that particular case, (5.5) implies that $\lambda=-1$ is the only nonzero eigenvalue.

Finally, we have given a formal proof that all the nonzero eigenvalues of (5.3) can be written as $\lambda = -\kappa + iku$, with $0<\kappa<1$ and $u\in\mathbb{R}$. The $L^2(\mathrm{d}x\mathrm{d}\gamma)$ norm of the solutions to this equation will then behave as

$$\Re e\exp((-\kappa+iku)t) = \exp(-\kappa t)\cos(k\,u\,t),$$

namely decay exponentially in time, while oscillating at a period proportional to $k=2\pi/R$. This is also in good agreement with the results from [40], where the spectrum of a fully discretized linear BGK equation is computed and shown to behave similarly.

To verify this interpretation experimentally, we compute in Table 2 the oscillation frequency ν with respect to the box size R. We observe that the product $R\nu$ appears indeed almost constant, yielding $u\simeq 0.35$ in the ongoing interpretation and giving more evidences that our numerical scheme preserves expected theoretical behaviors of the continuous equation.

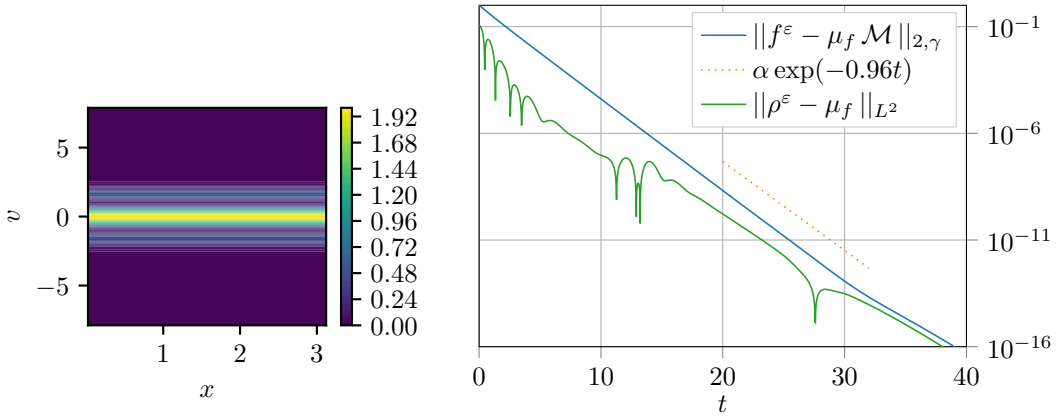


FIGURE 10. **Test 5.** Left: Equilibrium distribution ($t = 40$) in the $(x, v)-$phase-plane. Right: Time evolution of the weighted L^2 norm of the difference between f^ε and the global equilibrium, $\varepsilon = 1$, **Non-Gaussian BGK** case.

5.4. **The Non-Gaussian BGK case.**

Test 5. Relaxation towards a non-Gaussian equilibrium. As a final test case, we consider the relaxation of the solution to the linear BGK equation towards the following equilibrium distribution with polynomial tail:

$$M(v) = \frac{1}{1 + 0.1|v|^6} \left(\cos(\pi v) + 1.1\right). \tag{5.6}$$

This distribution verifies the hypothesis (3.1), namely that once projected on the velocity grid, it is of discrete mass 1, and has uniformly bounded discrete moments of order 2 and 4. According to Theorem 2, the solution to the discrete BGK equation solved with the scheme (3.10) will exhibit a discrete hypocoercive behavior.

We choose the very far from equilibrium initial condition (5.1) represented on the left side of Figure 10. Because of the fine structures in the equilibrium profile, we chose a more refined grid than in the previous test cases, namely $L = 35$, $N = 101$ and $\Delta t = 0.01$. We observe on the right of Figure 10 that the solution decays exponentially towards the non-Gaussian equilibrium (5.6).

6. Appendix. On discrete Poincaré inequalities

For the sake of completeness, we give in this Appendix the proofs of Poincaré inequalities we used in this article. We first start with the discrete Gaussian Poincaré inequality on bounded domain.

Proof of Lemma 5. We follow the same guidelines as in [22, Proposition 4.7]. Replacing if necessary f by $f - \rho$, where $\rho = \sum_{j\in\mathcal{J}} f_j \Delta v_j$, we can assume without loss of generality that $\rho = 0$.

First we can remark that

$$\|f\|_{2,\gamma}^2 = \sum_{j\in\mathcal{J}} (f_j\gamma_j)^2 \mathcal{M}_j \Delta v_j = \frac{1}{2} \sum_{(j,k)\in\mathcal{J}^2} (f_k\gamma_k - f_j\gamma_j)^2 \mathcal{M}_j\mathcal{M}_k\Delta v_j\Delta v_k$$
$$= \sum_{j<k} (f_k\gamma_k - f_j\gamma_j)^2 \mathcal{M}_j\mathcal{M}_k\Delta v_j\Delta v_k.$$

For $j<k$, we have

$$f_k\gamma_k - f_j\gamma_j = \sum_{l=j}^{k-1} (f_{l+1}\gamma_{l+1} - f_l\gamma_l).$$

Then, we obtain using Cauchy-Schwarz inequality

$$\|f\|_{2,\gamma}^2 = \sum_{j<k} \left(\sum_{l=j}^{k-1} \frac{f_{l+1}\gamma_{l+1} - f_l\gamma_l}{\sqrt{\Delta v_{l+\frac12}}} \sqrt{\Delta v_{l+\frac12}} \right)^2 \mathcal{M}_j\mathcal{M}_k\Delta v_j\Delta v_k$$
$$\leq \sum_{j<k} \left(\sum_{l=j}^{k-1} \left(D_v(f\gamma)_{l+\frac12} \right)^2 \Delta v_{l+\frac12} \right) (v_k - v_j)\mathcal{M}_j\mathcal{M}_k\Delta v_j\Delta v_k.$$

Let us introduce $F = (F_j)_{j\in\mathcal{J}}$ defined by

$$F_j := -\sum_{l=j}^{-1} \left(D_v(f\gamma)_{l+\frac12} \right)^2 \Delta v_{l+\frac12} \quad \text{if } j \leq -1,$$
$$F_j := \sum_{l=0}^{j-1} \left(D_v(f\gamma)_{l+\frac12} \right)^2 \Delta v_{l+\frac12} \quad \text{if } j \geq 0,$$

so that for $j<k$, we have

$$F_k - F_j = \sum_{l=j}^{k-1} \left(D_v(f\gamma)_{l+\frac12} \right)^2 \Delta v_{l+\frac12}.$$

It yields

$$\|f\|_{2,\gamma}^2 \leq \sum_{j<k} (F_k - F_j)(v_k - v_j)\mathcal{M}_j\mathcal{M}_k\Delta v_j\Delta v_k$$
$$\leq \frac12 \sum_{(j,k)\in\mathcal{J}^2} (F_k - F_j)(v_k - v_j)\mathcal{M}_j\mathcal{M}_k\Delta v_j\Delta v_k$$

since $(F_k - F_j)(v_k - v_j) = (F_j - F_k)(v_j - v_k)$.

Developing the right-hand side and using (3.2), we get

$$\|f\|_{2,\gamma}^2 \leq \sum_{j\in\mathcal{J}} F_j v_j \mathcal{M}_j \Delta v_j.$$

Now, using the definition of $\mathcal{M}_j$ given in (3.2), we obtain

$$\|f\|_{2,\gamma}^2 \leq \sum_{j\in\mathcal{J}} F_j (\mathcal{M}^*_{j-\frac12} - \mathcal{M}^*_{j+\frac12}).$$

Applying a discrete integration by parts, it gives

$$\|f\|_{2,\gamma}^2 \leq \sum_{j\in\mathcal{J}^*} (F_{j+1} - F_j)\mathcal{M}^*_{j+\frac12},$$

which finally yields the result since by definition,

$$F_{j+1} - F_j = \left(D_v(f\gamma)_{j+\frac{1}{2}}\right)^2 \Delta v_{j+\frac{1}{2}} = \left(D_v\left(\frac{f}{\mathcal{M}}\right)_{j+\frac{1}{2}}\right)^2 \Delta v_{j+\frac{1}{2}}.$$

□

Our numerical scheme and every result in the paper generalize in higher dimension on Cartesian meshes by tensorization. Hereafter we just show how it is done for one important point of our analysis: the discrete Gaussian Poincaré inequality.

Corollary 3 (Discrete Gaussian Poincaré inequality in higher dimension)**.** *Let $d \geq 1$. Given a discrete unidimensional Maxwellian $(\mathcal{M}_k)_{k\in\mathcal{J}}$ and $(\mathcal{M}^*_{k+1/2})_{k\in\mathcal{J}^*}$ satisfying* (3.2)*, a multi-index $j = (j_1, \ldots, j_d)$ and $e \in \{1, \ldots, d\}$, let us define multidimensional Maxwellians at cell centers and edge/face centers respectively by*

$$\mathcal{M}_j \;:=\; \prod_{e=1}^{d} \mathcal{M}_{j_e}\,, \quad \mathcal{M}_{j+\delta_e/2} \;:=\; \mathcal{M}_{j_e+1/2} \prod_{\substack{l=1\\ l\neq e}}^{d} \mathcal{M}_{j_l}\,,$$

where δ_e denotes the multi-index whose all components are zero, except the e-th. We also write $\Delta v_j \;:=\; \prod_{e=1}^d \Delta v_{j_e}$ and $\Delta v_{j+\delta_e/2} \;:=\; \Delta v_{j_e+1/2} \prod_{\substack{l=1\\ l\neq e}}^{d} \Delta v_{j_l}$. Then for all $f = (f_j)_{j\in\mathcal{J}^d} \in \mathbb{R}^{\mathcal{J}^d}$, one has that

$$\sum_{j\in\mathcal{J}^d} |f_j - \rho\mathcal{M}_j|^2\, \mathcal{M}_j^{-1}\, \Delta v_j \;\leq \tag{6.1}$$
$$\sum_{e=1}^{d} \sum_{j\in\mathcal{J}^{e-1}\times\mathcal{J}^*\times\mathcal{J}^{d-e}} \left(\frac{(f/\mathcal{M})_{(j_1,\ldots,j_e+1,\ldots,j_d)} - (f/\mathcal{M})_{(j_1,\ldots,j_e,\ldots,j_d)}}{\Delta v_{j+\delta_e/2}}\right)^2 \Delta v_{j+\delta_e/2}\, \mathcal{M}_{j+\delta_e/2}\,,$$

where $\rho = \sum_{j\in\mathcal{J}^d} f_j \Delta v_j$.

Proof. Let us introduce the notation $f_j^{(e)} = \sum_{(k_1,\ldots,k_e)\in\mathcal{J}^e} g_{(k_1,\ldots,k_e,j_{e+1},\ldots,j_d)} \prod_{l=1}^{e} \Delta v_{k_l}$. Observe that $f^{(0)} = f$ and that $f^{(d)} = \rho$. Then, use Pythagoras equality and apply $e-1$ times Jensen inequality (since $\sum_k \mathcal{M}_k \Delta v_k = 1$) to get

$$\begin{aligned}
&\textstyle\sum_{j\in\mathcal{J}^d} |f_j - \rho\mathcal{M}_j|^2\, \mathcal{M}_j^{-1}\, \Delta v_j \\
=\;& \textstyle\sum_{j\in\mathcal{J}^d} |\sum_{e=1}^{d} (f_j^{(e-1)} \prod_{u=1}^{e-1} \mathcal{M}_{j_u} - f_j^{(e)} \prod_{u=1}^{e} \mathcal{M}_{j_u})|^2\, \mathcal{M}_j^{-1}\, \Delta v_j \\
=\;& \textstyle\sum_{e=1}^{d} \sum_{j\in\mathcal{J}^d} |f_j^{(e-1)} \prod_{u=1}^{e-1} \mathcal{M}_{j_u} - f_j^{(e)} \prod_{u=1}^{e} \mathcal{M}_{j_u}|^2\, \mathcal{M}_j^{-1}\, \Delta v_j \\
=\;& \textstyle\sum_{e=1}^{d} \sum_{(j_e,\ldots,j_d)\in\mathcal{J}^{(d-e+1)}} |f_j^{(e-1)} - f_j^{(e)} \mathcal{M}_{j_e}|^2 \prod_{u=e}^{d} \mathcal{M}_{j_u}^{-1} \Delta v_{j_u} \\
\leq\;& \textstyle\sum_{e=1}^{d} \sum_{j\in\mathcal{J}^d} |f_j - \mathcal{M}_{j_e} \sum_{k_e\in\mathcal{J}} f_{(j_1,\ldots,j_{e-1},k_e,j_{e+1},\ldots,j_d)} \Delta v_{k_e}|^2\, \mathcal{M}_j^{-1}\, \Delta v_j
\end{aligned}$$

and one concludes by using d times the 1D discrete Gaussian Poincaré inequality (3.4). □

We finally provide a proof for the discrete Poincaré inequality on the (spatial) torus, in 1D and on a uniform mesh to simplify the presentation. The tensorization of the inequality may be done as before if one wants to generalize it to the multidimensional Cartesian setting.

Proof of Lemma 6. Let us assume that the grid is uniform in x, namely that $\Delta x_i = \Delta x > 0$, for any $i \in \mathcal{I} = \{0, \ldots, N-1\}$. We will also assume that the indices are $N-$periodic. This will allow to write a much simpler proof, using discrete Fourier transform (DFT). The result yet holds true in the general case, as long as N is **odd**.

Let us define the DFT of an $N-$periodic vector $g \in \mathbb{C}^N$, denoted by $\mathcal{F}_N g \in \mathbb{C}^N$, by:

$$(\mathcal{F}_N g)_k := \Delta x \sum_{l=0}^{N-1} g_l \, e^{-2i\pi kl\Delta x}, \quad \forall k \in \{0, \dots, N-1\}.$$

Such a discrete transform verifies the following discrete Parseval identity

$$\sum_{l=0}^{N-1} |g_l|^2 \Delta x = \sum_{k=0}^{N-1} \left|(\mathcal{F}_N g)_k\right|^2. \tag{6.2}$$

Since ϕ is of zero mass, one has that $(\mathcal{F}_N \phi)_0 = 0$. Then using identity (6.2), we have

$$\begin{aligned} \|D_x \phi\|_{L^2}^2 &= \sum_{l=0}^{N-1} \left|(D_x \phi)_l\right|^2 \Delta x \\ &= \sum_{k=0}^{N-1} \left|(\mathcal{F}_N D_x \, \phi)_k\right|^2 . \end{aligned} \tag{6.3}$$

Since $(D_x \phi)_l = (\phi_{l+1} - \phi_{l-1})/(2\Delta x)$, one has using the definition of the DFT that for any $k \in \{1, \dots, N-1\}$

$$\begin{aligned} (\mathcal{F}_N D_x \, \phi)_k &= -\frac{e^{2i\pi k\Delta x} - e^{-2i\pi k\Delta x}}{2\Delta x} \, (\mathcal{F}_N \, \phi)_k \\ &= -i \, \frac{\sin(2\pi k\Delta x)}{\Delta x} \, (\mathcal{F}_N \phi)_k . \end{aligned} \tag{6.4}$$

If N is even, then for $k = N/2$ the right hand side of (6.4) will be equal to 0. Since N is odd, and $(\mathcal{F}_N \phi)_0 = 0$, one has

$$\|D_x \phi\|_{L^2}^2 \geq S_*^2 \sum_{k=1}^{N-1} \left|(\mathcal{F}_N \phi)_k\right|^2 \tag{6.5}$$

where

$$\begin{aligned} S_* &:= \min_{k \in \{1, \dots, N-1\}} \left\{ \left| \frac{\sin(2\pi k\Delta x)}{\Delta x} \right| \right\} \\ &= \min \left\{ \left| \frac{\sin(2\pi\Delta x)}{\Delta x} \right|, \left| \frac{\sin(\pi(N-1)\Delta x)}{\Delta x} \right| \right\} \\ &= N \left| \sin\left(\frac{\pi}{N}\right) \right| \end{aligned} \tag{6.6}$$

because $\Delta x = 1/N$. Finally, using the Parseval identity (6.2) in (6.5) proves Lemma 6, with $C_P = 1/S_*$.

□

Remark 8. *In the continuous setting, the optimal constant for this Poincaré inequality is $C_P = 1/(2\pi)$. In our case, when $N \to \infty$, $S_* \to \pi$, which is suboptimal. This is due to the fact that D_x is a centered discretization of ∂_x. Taking a decentered discretization would yield $S_* \to 2\pi$.*

References

[1] Allaire, G., Blanc, X., Despres, B., and Golse, F. Transport et diffusion. Lecture notes, downloaded from `http://paestel.fr/content/transport-et-diffusion-g-allaire-x-blanc-b-despres-f-golse` in Nov. 2018.

[2] Bardos, C., Golse, F., Perthame, B., and Sentis, R. The nonaccretive radiative transfer equations: existence of solutions and Rosseland approximation. *J. Funct. Anal. 77*, 2 (1988), 434–460.

[3] Bardos, C., Santos, R., and Sentis, R. Diffusion approximation and computation of the critical size. *Trans. Amer. Math. Soc. 284*, 2 (1984), 617–649.
[4] Bensoussan, A., Lions, J. L., and Papanicolaou, G. C. Boundary layers and homogenization of transport processes. *Publications of the Research Institute for Mathematical Sciences 15*, 1 (1979), 53–157.
[5] Bessemoulin-Chatard, M., and Filbet, F. A finite volume scheme for nonlinear degenerate parabolic equations. *SIAM J. Sci. Comput. 34*, 5 (2012), B559–B583.
[6] Bessemoulin-Chatard, M., Herda, M., and Rey, T. Python implementation of an AP and hypocoercive numerical method for solving linear kinetic equations preserving mass. Binder-hosted Jupyter Notebook, available from `https://gitlab.com/thoma.rey/FV_HipoDiff`.
[7] Bhatnagar, P., Gross, E., and Krook, M. A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-component Systems. *Phys. Rev. 94*, 3 (1954), 511–525.
[8] Bouchut, F., Golse, F., and Pulvirenti, M. *Kinetic equations and asymptotic theory*, vol. 4 of *Series in Applied Mathematics (Paris)*. Gauthier-Villars, Éditions Scientifiques et Médicales Elsevier, Paris, 2000. Edited and with a foreword by Benoît Perthame and Laurent Desvillettes.
[9] Buet, C., and Dellacherie, S. On the Chang and Cooper scheme applied to a linear Fokker-Planck equation. *Commun. Math. Sci. 8*, 4 (2010), 1079–1090.
[10] Burger, M., Carrillo, J. A., and Wolfram, M.-T. A mixed finite element method for nonlinear diffusion equations. *Kinet. Relat. Models 3*, 1 (2010), 59–83.
[11] Cercignani, C., Illner, R., and Pulvirenti, M. *The Mathematical Theory of Dilute Gases*, vol. 106 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1994.
[12] Chainais-Hillairet, C., and Filbet, F. Asymptotic behaviour of a finite-volume scheme for the transient drift-diffusion model. *IMA J. Numer. Anal. 27*, 4 (2007), 689–716.
[13] Chainais-Hillairet, C., and Herda, M. Large-time behavior of a family of finite volume schemes for boundary-driven convection-diffusion equations. *arXiv preprint arXiv:1810.01087, to appear in IMA J. Numer. Anal.* (2019).
[14] Chang, J., and Cooper, G. A practical difference scheme for Fokker-Planck equations. *J. Comput. Phys. 6*, 1 (1970), 1–16.
[15] Crestetto, A., Crouseilles, N., and Lemou, M. A particle micro-macro decomposition based numerical scheme for collisional kinetic equations in the diffusion scaling. arXiv preprint 1701.05069.
[16] Crouseilles, N., Hivert, H., and Lemou, M. Numerical schemes for kinetic equations in the anomalous diffusion limit. Part I: the case of heavy-tailed equilibrium. *SIAM J. Sci. Comput. 38*, 2 (2016), A737–A764.
[17] Degond, P., Goudon, T., and Poupaud, F. Diffusion limit for nonhomogeneous and non-micro-reversible processes. *Indiana Univ. Math. J. 49*, 3 (2000), 1175–1198.
[18] Degond, P., and Mas-Gallic, S. Existence of solutions and diffusion approximation for a model Fokker-Planck equation. In *Proceedings of the conference on mathematical methods applied to kinetic equations (Paris, 1985)* (1987), vol. 16, pp. 589–636.
[19] Dimarco, G., and Pareschi, L. Numerical methods for kinetic equations. *Acta Num. 23* (2014), 369–520.
[20] Dimarco, G., Pareschi, L., and Samaey, G. Asymptotic-Preserving Monte Carlo methods for transport equations in the diffusive limit. *SIAM J. Sci. Comput. 40* (2018), A504–A528.
[21] Dolbeault, J., Mouhot, C., and Schmeiser, C. Hypocoercivity for linear kinetic equations conserving mass. *Trans. Amer. Math. Soc. 367*, 6 (2015), 3807–3828.
[22] Dujardin, G., Hérau, F., and Lafitte, P. Coercivity, hypocoercivity, exponential time decay and simulations for discrete Fokker-Planck equations. Preprint, https://hal.archives-ouvertes.fr/hal-01702545, 2018.
[23] Eymard, R., Gallouët, T., and Herbin, R. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII. North-Holland, Amsterdam, 2000, pp. 713–1020.
[24] Filbet, F., and Herda, M. A finite volume scheme for boundary-driven convection-diffusion equations with relative entropy structure. *Numer. Math. 137*, 3 (2017), 535–577.
[25] Filbet, F., Mouhot, C., and Pareschi, L. Solving the Boltzmann Equation in $Nlog_2N$. *SIAM J. Sci. Comput. 28*, 3 (2007), 1029–1053.
[26] Foster, E. L., Lohéac, J., and Tran, M. B. A structure preserving scheme for the Kolmogorov–Fokker–Planck equation. *Journal of Computational Physics 330* (2017), 319–339.
[27] Georgoulis, E. H. Hypocoercivity-compatible finite element methods for the long-time computation of Kolmogorov's equation. Preprint, https://arxiv.org/abs/1810.05137, 2018.
[28] Gosse, L., and Toscani, G. Identification of asymptotic decay to self-similarity for one-dimensional filtration equations. *SIAM J. Numer. Anal. 43*, 6 (2006), 2590–2606.
[29] Goudon, T., and Poupaud, F. Approximation by homogenization and diffusion of kinetic equations. *Comm. Partial Differential Equations 26*, 3-4 (2001), 537–569.
[30] Hérau, F. Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation. *Asymptot. Anal. 46*, 3-4 (2006), 349–359.

[31] Hérau, F. Introduction to hypocoercive methods and applications for simple linear inhomogeneous kinetic models. In *Lectures on the analysis of nonlinear partial differential equations. Part 5*, vol. 5 of *Morningside Lect. Math.* Int. Press, Somerville, MA, 2018, pp. 119–147.
[32] Hérau, F., and Nier, F. Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal. 171*, 2 (2004), 151–218.
[33] Hörmander, L. Hypoelliptic second order differential equations. *Acta Math. 119* (1967), 147–171.
[34] Il'In, A. A difference scheme for a differential equation with a small parameter multiplying the highest derivative. *Math. Zametki 6* (1969), 237–248.
[35] Jin, S. Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review. In *Lecture Notes for Summer School on "Methods and Models of KineticTheory"(M&MKT), Porto Ercole (Grosseto, Italy)* (2010), pp. 177–216.
[36] Jin, S., Pareschi, L., and Toscani, G. Uniformly accurate diffusive relaxation schemes for multiscale transport equations. *SIAM J. Numer. Anal. 38*, 3 (2000), 913–936.
[37] Jüngel, A. *Entropy methods for diffusive partial differential equations*. SpringerBriefs in Mathematics. Springer, [Cham], 2016.
[38] Kato, T. *Perturbation Theory for Linear Operators*. Springer, 1966.
[39] Klar, A. An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit. *SIAM J. Numer. Anal. 35*, 3 (1998), 1073–1094.
[40] Lafitte, P., Lejon, A., and Samaey, G. A high-order asymptotic-preserving scheme for kinetic equations using projective integration. *SIAM Journal on Numerical Analysis 54*, 1 (2016), 1–33.
[41] Larsen, E. W., and Keller, J. B. Asymptotic solution of neutron transport problems for small mean free paths. *J. Mathematical Phys. 15* (1974), 75–81.
[42] Lemou, M., and Mieussens, L. A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit. *SIAM J. Sci. Comput. 31*, 10 (2008), 334–368.
[43] Lions, P. L., and Toscani, G. Diffusive limit for finite velocity Boltzmann kinetic models. *Rev. Mat. Iberoamericana 13*, 3 (1997), 473–513.
[44] Liu, J.-G., and Mieussens, L. Analysis of an asymptotic preserving scheme for linear kinetic equations in the diffusion limit. *SIAM J. Numer. Anal. 48*, 4 (2010), 1474–1491.
[45] Mellet, A., Mischler, S., and Mouhot, C. Fractional diffusion limit for collisional kinetic equations. *Archive for Rational Mechanics and Analysis 199*, 2 (2011), 493–525.
[46] Mouhot, C., and Neumann, L. Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus. *Nonlinearity 19*, 4 (2006), 969–998.
[47] Pareschi, L., and Rey, T. Residual equilibrium schemes for time dependent partial differential equations. *Comput. & Fluids 156* (2017), 329–342.
[48] Porretta, A., and Zuazua, E. Numerical hypocoercivity for the Kolmogorov equation. *Math. Comp. 86*, 303 (2017), 97–119.
[49] Poupaud, F. Diffusion approximation of the linear semiconductor Boltzmann equation: analysis of boundary layers. *Asymptot. Anal. 4*, 4 (1991), 293–317.
[50] Scharfetter, D., and Gummel, H. Large signal analysis of a silicon Read diode. *IEEE Trans. Elec. Dev. 16* (1969), 64–77.
[51] Schmeiser, C. Entropy methods. `https://homepage.univie.ac.at/christian.schmeiser/Entropy-course.pdf`, 2018.
[52] Schmeiser, C., and Zwirchmayr, A. Convergence of moment methods for linear kinetic equations. *SIAM J. Numer. Anal. 36*, 1 (1999), 74–88.
[53] Villani, C. Hypocoercivity. *Mem. Amer. Math. Soc. 202* (2009), 184.

Marianne Bessemoulin-Chatard
Université de Nantes, Laboratoire de Mathématiques Jean Leray, UMR CNRS 6629
F-44000 Nantes, France
Email: marianne.bessemoulin@univ-nantes.fr

Maxime Herda
Inria, Univ. Lille, CNRS, UMR 8524 – Laboratoire Paul Painlevé
F-59000 Lille, France

Email: maxime.herda@inria.fr

Thomas Rey
Univ. Lille, CNRS, UMR 8524, Inria – Laboratoire Paul Painlevé
F-59000 Lille, France
Email: thomas.rey@univ-lille.fr