Regularity of velocity averages for transport equations on random discrete velocity grids

Nathalie Ayi[1,2] and Thierry Goudon[1,2]

[1]Project COFFEE, INRIA Sophia Antipolis Méditerranée
[2] Labo J.-A. Dieudonné, UMR 7351 CNRS-Univ. Nice Sophia Antipolis

Abstract

We go back to the question of the regularity of the "velocity average" $\int f(x,v)\psi(v)\,\mathrm{d}\mu(v)$ when f and $v\cdot\nabla_x f$ both belong to L^2, and the variable v lies in a discrete subset of $\mathbb{R}^D$. First of all, we provide a rate, depending on the number of velocities, to the defect of $H^{1/2}$ regularity. Second of all, we show that the $H^{1/2}$ regularity holds in expectation when the set of velocities is chosen randomly. We apply this statement to investigate the consistency with the diffusion asymptotics of a Monte–Carlo–like discrete velocity model.

1 Introduction

Averaging lemma is now a classical tool for the analysis of kinetic equations. Roughly speaking it can be explained as follows. Let $\mathscr{V}\subset\mathbb{R}^D$, endowed with a measure $\mathrm{d}\mu$. We consider a sequence of functions $f_n:\mathbb{R}^D\times\mathscr{V}\to\mathbb{R}$. We assume that

a) $\big(f_n\big)_{n\in\mathbb{N}}$ is bounded in $L^2(\mathbb{R}^D\times\mathscr{V})$,

b) $\big(v\cdot\nabla_x f_n\big)_{n\in\mathbb{N}}$ is bounded in $L^2(\mathbb{R}^D\times\mathscr{V})$.

Given $\psi\in C_c^\infty(\mathbb{R}^D)$, we are interested in the *velocity average*

$$\rho_n[\psi](x)=\int_{\mathscr{V}} f_n(x,v)\psi(v)\,\mathrm{d}\mu(v).$$

Of course, a) already tells us that $\big(\rho_n[\psi]\big)_{n\in\mathbb{N}}$ is bounded in $L^2(\mathbb{R}^D)$. We wish to obtain further regularity or compactness properties, as a consequence of the additional assumption b), and the fact that we are averaging with respect to the variable v. The first result in that direction dates back to [3] (see also [1]); it asserts that $\big(\rho_n[\psi]\big)_{n\in\mathbb{N}}$ is bounded in the Sobolev space $H^{1/2}(\mathbb{R}^D)$ and it is thus relatively compact in $L^2_{\mathrm{loc}}(\mathbb{R}^D)$, by virtue of the standard Rellich's theorem. This basic result has been improved in many directions: L^2 can be replaced by the L^p framework, at least with $1<p<\infty$, and we can relax b) by allowing derivatives with respect to v and certain loss of regularity with respect to x; see, among others, [13, 16, 26]. Time derivative or force terms can be considered as well, see, additionally to the above-mentionned references, [5]. Such an argument plays a crucial role in the stunning theory of "renormalized solutions" of the Boltzmann equation [12], and more generally for proving the existence of solutions to non linear kinetic models like in [11]. It is equally a crucial ingredient for the analysis of hydrodynamic regimes, which establish the connection between microscopic models and fluid mechanics systems; for the asymptotic of the Boltzmann equation to the incompressible Navier–Stokes system, which needs a suitable L^1 version of the average lemma [17], we refer the reader to [18, 27, 29]. Finally, it is worth pointing that averaging lemma can be used to investigate the regularizing effects of certain PDE (convection-diffusion and elliptic equations, nonlinear conservation laws, etc) [28].

In order to illustrate our purpose, let us consider the following simple model which can be motivated from radiative transfer theory:

$$\varepsilon\partial_t f_\varepsilon+v\cdot\nabla_x f_\varepsilon=\frac{1}{\varepsilon}\sigma(\rho_\varepsilon)(\rho_\varepsilon-f_\varepsilon) \tag{1.1}$$

where

$$\rho_\varepsilon(t,x) = \int_{\mathscr{V}} f_\varepsilon(t,x,v)\,\mathrm{d}\mu(v),$$

and $\sigma : [0,\infty) \to [0,\infty)$ is a given smooth function. The parameter $0 < \varepsilon \ll 1$ is defined from physical quantities. As it tends to 0, both $f_\varepsilon(t,x,v)$ and $\rho_\varepsilon(t,x)$ converge to $\rho(t,x)$, which satisfies the non linear diffusion equation

$$\partial_t \rho = \nabla_x \cdot \Big(A \nabla_x F(\rho)\Big), \quad A = \int_{\mathscr{V}} v \otimes v\,\mathrm{d}\mu(v), \quad F(\rho) = \int_0^\rho \frac{\mathrm{d}z}{\sigma(z)}. \tag{1.2}$$

The averaging lemma is an efficient tool to deal with the nonlinearity of such a problem, as discussed in [3].

However the discussion above hides the fact that we need some assumptions on the measured set of velocities $(\mathscr{V}, \mathrm{d}\mu)$ in order to obtain the regularization property of the velocity averaging. Roughly speaking, we need "enough" directions v when we consider the derivatives in b). More technically, the compactness statement holds provided for any $0 < R < \infty$ we can find $C_R > 0$, $\delta_0 > 0$, $\gamma > 0$ such that for $0 < \delta < \delta_0$ and $\xi \in \mathbb{S}^{N-1}$, we have

$$\mathrm{meas}\Big(\Big\{v \in \mathscr{V} \cap B(0,R),\ |v \cdot \xi| \le \delta\Big\}\Big) \le C_R \delta^\gamma.$$

This assumption appears in many statements about regularity of the velocity averages; as far as we are only interested in compactness issue, it can be replaced by the more intuitive assumption (see e. g. [14, Th. 1 in Lect. 3]): for any $\xi \in \mathbb{S}^{N-1}$ we have

$$\mathrm{meas}\Big(\Big\{v \in \mathscr{V} \cap B(0,R),\ v \cdot \xi = 0\Big\}\Big) = 0. \tag{1.3}$$

Clearly these assumptions are satisfied when the measure $\mathrm{d}\mu$ is absolutely continuous with respect to the Lebesgue measure (with, for the sake of concreteness, $\mathscr{V} = \mathbb{R}^D$ or $\mathscr{V} = \mathbb{S}^{D-1}$). However, they fail for models based on a discrete set of velocities. For instance let $\mathscr{V} = \{v_1, ..., v_N\}$, with $v_j \in \mathbb{R}^D$, and $\mathrm{d}\mu(v) = \frac{1}{N}\sum_{j=1}^N \delta(v = v_j)$; it suffices to pick $\xi \in \mathbb{S}^{N-1}$ orthogonal to one of the v_j's to contradicts (1.3). (Note that alternative proofs based on compensated compactness techniques have been proposed to justify the asymptotic regime from (1.1) to (1.2), that apply to certain discrete velocity models, see [10, 20, 23].) Nevertheless, when the discrete velocities come from a discretization grid of the whole space, the averaging lemma can be recovered asymptotically letting the mesh step go to 0, as shown in [25], motivated by the convergence analysis of numerical schemes for the Boltzmann equation.

This paper aims at investigating further these issues. To be more specific, in Section 2 we revisit the averaging lemma for discrete velocities in two directions. First of all, we make more precise the analysis of [25], obtaining a rate on the defect to the $H^{1/2}$ regularity of the velocity average, depending on the mesh size. Second of all, we establish a stochastic version of the averaging lemma. We are still working with a finite number of velocities on bounded sets; however, choosing the velocities randomly, the "compactifying" property of assumption b) can be restored by dealing with the expectation of $\rho_n[\psi]$. This is a natural way to involve "enough velocities", by looking at a large set of realization of the discrete velocity grid. The analysis is completed in Section 3 by going back to the asymptotic problem $\varepsilon \to 0$ in (1.1), with a random discretization of the velocity variable, in the spirit of the Monte–Carlo approach.

2 Discrete Velocity Averaging Lemmas

2.1 Deterministic case: evaluation of the defect

As mentioned above, it is a well known fact that, in the deterministic context, the averaging lemma fails for discrete velocity models. However, as mentioned by S. Mischler in [25], the compactness of velocity averages is recovered asymptotically when we refine a velocity grid in order to recover a continuous velocity model. Here, we wish to quantify the defect of compactness when the number of velocities is finite and fixed. This is the aim of the following claim which shows that the macroscopic density $\rho[\psi]$ "belongs to $H^{1/2}(\mathbb{R}^D) + O(\frac{1}{\sqrt{N}})L^2(\mathbb{R}^D)$".

Proposition 2.1. *Let $N \in \mathbb{N} \setminus \{0\}$ and define*

$$A_N = \left(\frac{1}{N}\mathbb{Z}\right)^D \cap [-0.5, 0.5]^D.$$

Let $f, g \in L^2(\mathbb{R}^D \times A_N)$ satisfy for all $k \in \mathbb{Z}$,

$$v_k \cdot \nabla_x f(x, v_k) = g(x, v_k). \tag{2.1}$$

We suppose that the L^2 norm of f and g is bounded uniformly with respect to N. Then, for all $\psi \in C_c^\infty(\mathbb{R}^D)$, the macroscopic quantity

$$\rho[\psi](x) = \frac{1}{(N+1)^D} \sum_k f(x, v_k)\psi(v_k)$$

can be split as $\rho[\psi](x) = \Theta[\psi](x) + \frac{1}{\sqrt{N}}\widetilde{\Delta[\psi]}(x)$ where $\Theta[\psi]$ and $\widetilde{\Delta[\psi]}$ are bounded uniformly with respect to N in $H^{1/2}(\mathbb{R}^D)$ and $L^2(\mathbb{R}^D)$ respectively.

Remark 2.1. *Note that in this statement N is the number of grid points per axis. Accordingly, there are $\mathcal{N} = (N+1)^D$ velocities in the set A_N. Therefore the defect of $H^{1/2}$ regularity decays like $\mathcal{N}^{1/2D}$, depending on the dimension.*

Proof. As usual, we start by applying the Fourier transform to (2.1). Then for all $k \in \mathbb{Z}$ and $\xi \in \mathbb{R}^D$, we get

$$\xi \cdot v_k \ \hat{f}(\xi, v_k) = (-i)\hat{g}(\xi, v_k).$$

Let us set

$$F(\xi) := \left(\frac{1}{(N+1)^D} \sum_k |\hat{f}(\xi, v_k)|^2\right)^{1/2}, \qquad G(\xi) := \left(\frac{1}{(N+1)^D} \sum_k |\hat{g}(\xi, v_k)|^2\right)^{1/2}.$$

By assumption, we have $F, G \in L^2_\xi$. Still following the standard arguments, we pick $\delta > 0$ and we split

$$\begin{aligned}
\hat{\rho}[\psi](\xi) &= \frac{1}{(N+1)^D} \sum_k \hat{f}(\xi, v_k)\psi(v_k) \\
&= \frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| < \delta|\xi|} \hat{f}(\xi, v_k)\psi(v_k) + \frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| \geq \delta|\xi|} \hat{f}(\xi, v_k)\psi(v_k).
\end{aligned}$$

The Cauchy–Schwarz inequality permits us to dominate the first term as follows

$$\begin{aligned}
&\left|\frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| < \delta|\xi|} \hat{f}(\xi, v_k)\psi(v_k)\right| \\
&\qquad \leq \|\psi\|_\infty \left(\frac{1}{(N+1)^D} \sum_k |\hat{f}(\xi, v_k)|^2\right)^{1/2} \left(\frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| < \delta|\xi|} 1\right)^{1/2}.
\end{aligned} \tag{2.2}$$

For the second term, we use the information in (2.1); it yields

$$\begin{aligned}
\left|\frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| \geq \delta|\xi|} \hat{f}(\xi, v_k)\psi(v_k)\right| &= \left|\frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| \geq \delta|\xi|} \frac{(-i)\hat{g}(\xi, v_k)}{\xi \cdot v_k} \ \psi(v_k)\right| \\
&\leq \|\psi\|_\infty \left(\frac{1}{(N+1)^D} \sum_k |\hat{g}(\xi, v_k)|^2\right)^{1/2} \left(\frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| \geq \delta|\xi|} \frac{1}{|\xi \cdot v_k|^2}\right)^{1/2}.
\end{aligned} \tag{2.3}$$

From now on we assume $\xi \neq 0$. Let $(e_1, .., , e_D)$ stand for the canonical basis of $\mathbb{R}^D$ so that $\xi = \sum_{j=1}^D \alpha_j e_j$ with $\alpha_j \in \mathbb{R}$. We distinguish the following two cases:

i) either ξ is aligned with an axis: all α_j's vanish but one,

ii) or ξ is generated by at least two vectors of the basis.

We start with the case i) assuming for instance $\xi = \alpha e_1$. Then $\xi \cdot v_k = \alpha v_k^1$, where v_k^1 is the first component of the vector v_k.

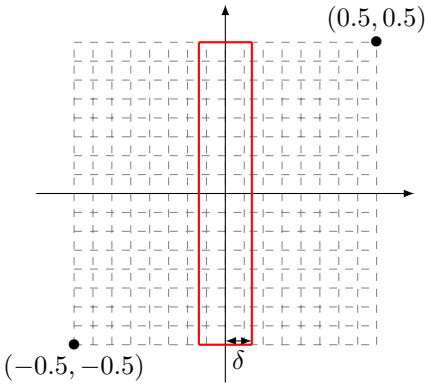

Figure 1: The delimited area corresponds to $|\xi \cdot v_k| < \delta|\xi|$ for ξ colinear to e_1.

We refer the reader to Fig. 1 for completing the discussion. On each horizontal line we find $2\lfloor \delta N \rfloor + 1$ velocities such that $|\xi \cdot v_k| < \delta|\xi|$, where $\lfloor s \rfloor$ stands for the integer part of s. Thus, since there is $(N+1)^{D-1}$ such lines on the domain A_N, we obtain

$$\sum_{|\xi \cdot v_k| < \delta|\xi|} 1 \quad = \quad (2\lfloor \delta N \rfloor + 1)(N+1)^{D-1} = 2\left(\delta + \frac{1}{N}\right)(N+1)^D.$$

Coming back to (2.2), we arrive at

$$\left| \frac{1}{(N+1)^D} \sum_{|\xi \cdot v_k| < \delta|\xi|} \hat{f}(\xi, v_k)\psi(v_k) \right| \leq \; C \; F(\xi)\sqrt{\delta + \frac{1}{N}}.$$

where $C > 0$ is a generic constant which does not depend on N and ξ.

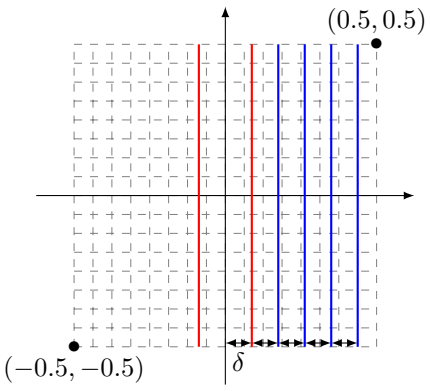

Figure 2: Splitting of the velocity space in strips of width δ. This space being symmetric, we only deal with the part corresponding to positive abscissae.

Next, we cover the set of velocities such that $|v_k \cdot \xi| \geq \delta|\xi|$ by strips of width δ, see Fig. 2 in dimension $D = 2$. We denote by S_p the p-th strip delimited by the straight lines $x = p\delta$ and $x = (p+1)\delta$. Each velocity on the strip S_p satisfies $p\delta \leq v_k^1 \leq (p+1)\delta$. Moreover, given a strip S_p, we cannot find more than $\lfloor \delta N \rfloor + 1$ abscissae in the strip and there is $(N+1)^{D-1}$

lines in the domain. It follows that

$$\begin{aligned}
\sum_{|\xi\cdot v_k|\geq\delta|\xi|}\frac{1}{|\xi\cdot v_k|^2} &= \sum_{|\xi\cdot v_k|\geq\delta|\xi|}\frac{1}{|\xi|^2}\frac{1}{|\frac{\xi}{|\xi|}.v_k|^2}\\
&\leq \frac{1}{|\xi|^2}2\left(\sum_{p\geq 1}\frac{1}{(p\delta)^2}\right)(\delta N+1)(N+1)^{D-1}\\
&\leq \frac{1}{|\xi|^2}2\left(\sum_{p\geq 1}\frac{1}{p^2}\right)\frac{1}{\delta}\left(1+\frac{1}{\delta N}\right)(N+1)^{D}.
\end{aligned}$$

Thus, we deduce from (2.3) that

$$\left|\frac{1}{(N+1)^D}\sum_{|\xi\cdot v_k|\geq\delta|\xi|}\hat{f}(\xi,v_k)\psi(v_k)\right|\leq C\ G(\xi)\frac{1}{|\xi|\sqrt{\delta}}\left(1+\frac{1}{\delta N}\right)^{1/2}.$$

We conclude that

$$|\hat{\rho}[\psi](\xi)| \quad\leq\quad C\left(F(\xi)\sqrt{\delta+\frac{1}{N}}+G(\xi)\frac{1}{|\xi|\sqrt{\delta}}\left(1+\frac{1}{\delta N}\right)^{1/2}\right). \tag{2.4}$$

holds when ξ is aligned to the axis.

We turn to the general case ii). As illustrated in Fig. 3, we can assume that the angle θ between ξ and one of the axis (say e_1) lies in $]0,\pi/4[$, the other cases follow by a symmetry argument.

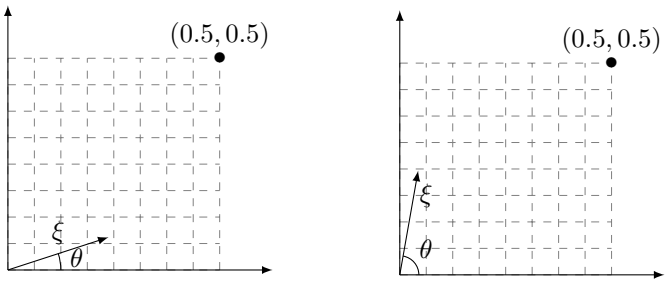

Figure 3: Representation of $\xi\in\mathbb{R}^2$ with $\theta\in]0,\frac{\pi}{4}]$ and $\theta\in]\frac{\pi}{4},\frac{\pi}{2}]$ with $\cos(\theta)|\xi|=\xi\cdot e_1$.

The reasoning still consists in counting velocities in strips appropriately defined. As said above, without loss of generality we can assume that $\theta\in]0,\frac{\pi}{4}]$, where we have set $\cos(\theta)|\xi|=\xi\cdot e_1$. We set $\ell_1:=\frac{\delta}{\cos\theta}$. On a given strip, we can find at most $(\lfloor\ell_1 N\rfloor+1)\times(N+1)^{D-1}$ velocities, see Fig. 5.

Therefore, bearing in mind that $0<\theta<\frac{\pi}{4}$, we obtain

$$\begin{aligned}
\sum_{|\xi\cdot v_k|\geq\delta|\xi|}\frac{1}{|\xi\cdot v_k|^2} &= \sum_{|\xi\cdot v_k|\geq\delta|\xi|}\frac{1}{|\xi|^2}\frac{1}{|\frac{\xi}{|\xi|}\cdot v_k|^2}\\
&\leq \frac{1}{|\xi|^2}\times 2\sum_{p\geq 1}\frac{1}{(p\delta)^2}\Big(\frac{\delta}{\cos\theta}N+1\Big)(N+1)^{D-1}\\
&\leq \frac{1}{|\xi|^2}\times 2\sum_{p\geq 1}\frac{1}{(p\delta)^2}\frac{1}{\delta\cos\theta}\left(1+\frac{1}{\delta N}\right)(N+1)^{D}\\
&\leq 2\sqrt{2}\frac{1}{|\xi|^2}\frac{1}{\delta}\left(1+\frac{1}{\delta N}\right)(N+1)^{D}
\end{aligned}$$

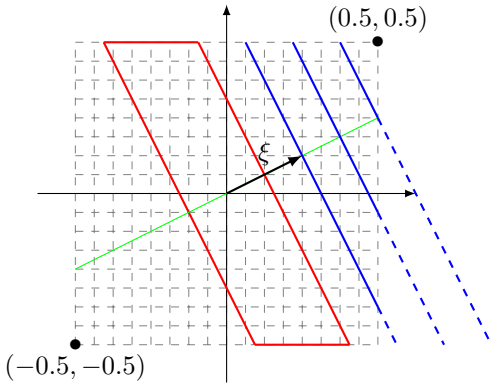

Figure 4: The area corresponding to $|\xi \cdot v_k| \leq \delta|\xi|$ is delimited as previously. The complementary set is split into strips of width δ.

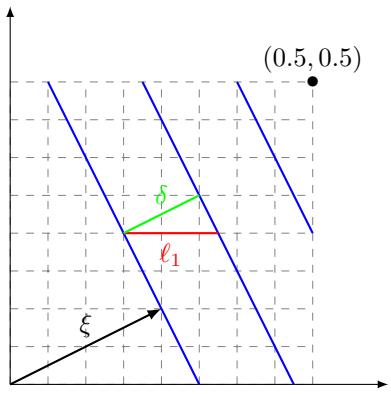

Figure 5: Representation of the parameter ℓ_1.

and

$$\begin{aligned}\sum_{|\xi\cdot v_k|<\delta|\xi|} 1 &= (2\lfloor \ell_1 N\rfloor + 1)(N+1)^{D-1} \\ &\leq 2\left(\frac{\delta}{\cos\theta}N + 1\right)(N+1)^{D-1} \\ &\leq 2\sqrt{2}\left(\delta + \frac{1}{N}\right)(N+1)^{D}.\end{aligned}$$

Thus, we deduce exactly like in case i) that (2.4) holds for any $\xi \neq 0$.

Therefore, we have established that for all $\xi \neq 0$, we get (2.4) for all $\delta > 0$. We take

$$\delta = \frac{1}{|\xi|}\mathbf{1}_{\{N\geq|\xi|\}} + \frac{1}{N}\mathbf{1}_{\{N<|\xi|\}}$$

and we denote

$$\Theta_N(\xi) := \hat{\rho}[\psi](\xi)\mathbf{1}_{\{N\geq|\xi|\}}, \qquad \Delta_N(\xi) := \hat{\rho}[\psi](\xi)\mathbf{1}_{\{N<|\xi|\}}.$$

Then, we have

$$\begin{aligned}\Theta_N(\xi) &\leq C\left(F(\xi)\sqrt{\frac{1}{|\xi|}+\frac{1}{N}}+G(\xi)\frac{1}{|\xi|\sqrt{\frac{1}{|\xi|}}}\left(1+\frac{1}{\frac{N}{|\xi|}}\right)^{1/2}\right)\mathbf{1}_{\{N\geq|\xi|\}}\\ &\leq C\left(F(\xi)+G(\xi)\right)\frac{1}{\sqrt{|\xi|}}.\end{aligned}$$

It implies that

$$|\xi|\Theta_N(\xi)^2 \leq C(G^2(\xi)+F^2(\xi)),$$

which equally holds true for $\xi=0$. Then by assumption on f and g, we deduce that $\Theta_N \in H^{1/2}(\mathbb{R}^D)$. Finally, we evaluate the remainder as follows

$$\begin{aligned}\Delta_N(\xi) &\leq C\left(F(\xi)\sqrt{\frac{2}{N}}+G(\xi)\frac{1}{|\xi|\sqrt{\frac{1}{N}}}\left(1+\frac{1}{\frac{1}{N}N}\right)\right)\mathbf{1}_{\{N<|\xi|\}}\\ &\leq \frac{C}{\sqrt{N}}(F(\xi)+G(\xi)).\end{aligned}$$

We conclude that

$$\Delta_N^2(\xi) \leq \frac{C}{N}\left(F^2(\xi)+G^2(\xi)\right),$$

which is also satisfied when $\xi=0$. Thus, by assumption on f and g, $\|\Delta_N\|_{L^2}$ is dominated by $\frac{1}{\sqrt{N}}$, an observation which finishes the proof. □

2.2 A stochastic discrete velocity averaging lemma

Dealing with random discrete velocities we can expect to make the defect vanish when taking the expectation of the velocity averages. This is indeed the case as shown in the following statement.

Theorem 2.2. *Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space. Let $V_1,...,V_{\mathscr{N}}$ be i.i.d. random variables, distributed according to the continuous uniform distribution on $[-0.5,0.5]^D$. We set*

$$\mathrm{d}\mu=\frac{1}{\mathscr{N}}\sum_{k=1}^{\mathscr{N}}\delta(v=V_k).$$

Let $f,g\in L^2(\mathbb{R}^D\times\mathbb{R}^D\times\Omega,\,\mathrm{d}x\,\mathrm{d}\mu(v)\,\mathrm{d}\mathbb{P})$ satisfy for all $x\in\mathbb{R}^D$, $\omega\in\Omega$, and $k\in\{1,...,\mathscr{N}\}$

$$V_k\cdot\nabla_x f(x,V_k)=g(x,V_k). \tag{2.5}$$

Then, for all $\psi\in C_c^\infty(\mathbb{R}^D)$, the macroscopic quantity

$$\rho[\psi](x):=\frac{1}{\mathscr{N}}\sum_{k=1}^{\mathscr{N}}f(x,V_k)\psi(V_k)=\int_{\mathbb{R}^D}f(x,v)\psi(v)\,\mathrm{d}\mu(v)$$

satisfies $\mathbb{E}\rho[\psi]\in H^{1/2}(\mathbb{R}^D)$ (and it is bounded in this space if the L^2 norm of f and g is bounded uniformly with respect to $\mathscr{N}$).

Remark 2.2. *We point out that this statement has a different nature form the stochastic averaging lemma devised in [8, 9], where the velocity set still satisfies an assumption like (1.3) but the equation for $v\cdot\nabla_x f_n$ involves a stochastic term. Our analysis is closer in spirit to the results in [22] where the velocity variable is deterministic but is is multiplied by a Brownian motion.*

Proof. We apply the Fourier transform to (2.5). Then, for all k, we get

$$\xi\cdot V_k\hat{f}(\xi,V_k)=(-i)\hat{g}(\xi,V_k).$$

We set

$$F(\xi):=\left(\frac{1}{\mathscr{N}}\mathbb{E}\sum_k|\hat{f}(\xi,V_k)|^2\right)^{1/2},\qquad G(\xi):=\left(\frac{1}{\mathscr{N}}\mathbb{E}\sum_k|\hat{g}(\xi,V_k)|^2\right)^{1/2}.$$

Let us split

$$\begin{aligned}\mathbb{E}\hat{\rho}[\psi](\xi) &= \mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_k \hat{f}(\xi,V_k)\psi(V_k)\right]\\ &= \mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_{|\xi\cdot V_k|<\delta|\xi|}\hat{f}(\xi,V_k)\psi(V_k)\right]+\mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_{|\xi\cdot V_k|\geq\delta|\xi|}\hat{f}(\xi,V_k)\psi(V_k)\right].\end{aligned}$$

for $\delta>0$. The Cauchy–Schwarz inequality leads to the following estimate: on the one hand

$$\begin{aligned}&\left|\mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_{|\xi\cdot V_k|<\delta|\xi|}\hat{f}(\xi,V_k)\psi(V_k)\right]\right|\\ &\qquad\leq\|\psi\|_\infty\left(\frac{1}{\mathscr{N}}\mathbb{E}\sum_k|\hat{f}(\xi,V_k)|^2\right)^{1/2}\left(\frac{1}{\mathscr{N}}\mathbb{E}\sum_{|\xi\cdot V_k|<\delta|\xi|}1\right)^{1/2},\end{aligned}$$

and, on the other hand

$$\begin{aligned}\left|\mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_{|\xi\cdot v_k|\geq\delta|\xi|}\hat{f}(\xi,V_k)\psi(V_k)\right]\right| &= \left|\mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_{|\xi\cdot V_k|\geq\delta|\xi|}\frac{(-i)\hat{g}(\xi,V_k)}{\xi\cdot V_k}\ \psi(V_k)\right]\right|\\ &\leq\|\psi\|_\infty\left(\frac{1}{\mathscr{N}}\mathbb{E}\sum_k|\hat{g}(\xi,V_k)|^2\right)^{1/2}\left(\frac{1}{\mathscr{N}}\mathbb{E}\sum_{|\xi\cdot v_k|\geq\delta|\xi|}\frac{1}{|\xi\cdot V_k|^2}\right)^{1/2}.\end{aligned}$$

We only detail the case where $\xi=\alpha e_1$, $\alpha\in\mathbb{R}$, the other cases being deduced by adapting the reasoning of the proof of Proposition 2.1. We have

$$\begin{aligned}\mathbb{E}\left[\sum_{|\xi\cdot V_k|\geq\delta|\xi|}\frac{1}{|\xi\cdot V_k|^2}\right] &= \mathbb{E}\left[\sum_{|\xi\cdot V_k|\geq\delta|\xi|}\frac{1}{|\xi|^2}\frac{1}{|\frac{\xi}{|\xi|}\cdot V_k|^2}\right]\\ &\leq \mathbb{E}\left[\frac{1}{|\xi|^2}2\left(\sum_{p\geq1}\frac{1}{(p\delta)^2}\right)M_p\right]\end{aligned}$$

where M_p is the number of velocities in the p-th strip (see Fig. 2). We bear in mind that M_p is a random variable: since the V_i's are distributed according to the uniform law, we have

$$\mathbb{P}(V_i\in S_p)=\delta$$

and, the variables $V_1,\ldots,V_{\mathscr{N}}$ being independent, M_p follows a binomial distribution of parameters $\mathscr{N}$ and δ. Therefore, we are led to

$$\begin{aligned}\mathbb{E}\left[\sum_{|\xi\cdot V_k|\geq\delta|\xi|}\frac{1}{|\xi\cdot V_k|^2}\right] &\leq \frac{1}{|\xi|^2}2\left(\sum_{p\geq1}\frac{1}{(p\delta)^2}\right)\mathbb{E}\left[M_p\right]\\ &\leq C\frac{1}{|\xi|^2\delta}\mathscr{N},\end{aligned}\tag{2.6}$$

which yields

$$\left|\mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_{|\xi\cdot V_k|\geq\delta|\xi|}\hat{f}(\xi,V_k)\psi(V_k)\right]\right|\leq CG(\xi)\frac{1}{|\xi|\sqrt{\delta}}.$$

On the same token, we get

$$\mathbb{E}\left[\sum_{|\xi\cdot V_k|<\delta|\xi|}1\right]=2\delta\mathscr{N}\tag{2.7}$$

so that

$$\left|\mathbb{E}\left[\frac{1}{\mathscr{N}}\sum_{|\xi\cdot V_k|<\delta|\xi|}\hat{f}(\xi,V_k)\psi(V_k)\right]\right|\leq CF(\xi)\sqrt{\delta}.$$

Finally, we arrive at

$$|\mathbb{E}\hat{\rho}[\psi](\xi)| \leq C\left(F(\xi)\sqrt{\delta} + \frac{G(\xi)}{|\xi|\sqrt{\delta}}\right).$$

We apply this inequality with $\delta = \frac{G(\xi)}{|\xi|F(\xi)}$, which leads to

$$|\mathbb{E}\hat{\rho}[\psi](\xi)| \leq C\sqrt{F(\xi)G(\xi)}\frac{1}{\sqrt{|\xi|}}.$$

It concludes the proof by using the assumptions on f and g. □

Remark 2.3. *We can readily extend the result to non uniform laws: we assume that the V_i's are identically and independently distributed in $\mathbb{R}^D$ according to a continuous and bounded density of probability Φ. The number M_p of velocities in the strip S_p still follows a binomial law but now the expectation value depends on Φ and M_p can be shown to be dominated by $\mathscr{N}\|\Phi\|_\infty\delta$.*

For certain applications, the variable v lies on the sphere. This is the case for the kinetic models arising in radiative transfer theory where v represents the *direction* of flight of photons, which, of course, all travel with the speed of light. We can adapt the stochastic averaging lemma to this situation.

Theorem 2.3. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $V_1, ..., V_{\mathscr{N}}$ be i.i.d. random variables, distributed according to the continuous uniform distribution on $\mathbb{S}^{D-1}$. We set*

$$\mathrm{d}\mu = \frac{1}{\mathscr{N}}\sum_{k=1}^{\mathscr{N}} \delta(v = V_k).$$

Let $f, g \in L^2(\mathbb{R}^D \times \mathbb{R}^D \times \Omega,\, \mathrm{d}x\,\mathrm{d}\mu(v)\,\mathrm{d}\mathbb{P})$ satisfy for all $x \in \mathbb{R}^D$, $\omega \in \Omega$, and $k \in \{1, ..., \mathscr{N}\}$

$$V_k \cdot \nabla_x f(x, V_k) = g(x, V_k).$$

Then, for all $\psi \in C_c^\infty(\mathbb{S}^{D-1})$, the macroscopic quantity

$$\rho[\psi](x) := \frac{1}{\mathscr{N}}\sum_{k=1}^{\mathscr{N}} f(x, V_k)\psi(V_k) = \int_{\mathbb{R}^D} f(x, v)\psi(v)\,\mathrm{d}\mu(v)$$

satisfies $\mathbb{E}\rho[\psi] \in H^{1/2}(\mathbb{R}^D)$.

Proof. The proof follows the same arguments as for Theorem 2.2; we only indicate the main changes. The proof still relies in counting the velocities produced by the random sampling in the domain

$$S_p = \left\{v \in \mathbb{S}^{D-1},\ \delta p|\xi| \leq |v \cdot \xi| \leq \delta(p+1)|\xi|\right\},$$

for given $\xi \in \mathbb{R}^D \setminus \{0\}$, $\delta > 0$ and $p \in \mathbb{Z}$. We define $\theta \in [0, 2\pi]$ such that

$$v \cdot \frac{\xi}{|\xi|} = \cos(\theta) \in [-1, +1].$$

Considering the random vectors V_k, the associated variable θ_k is randomly distributed on $[0, 2\pi]$. For symmetry reasons, $\mathbb{P}(V_k \in S_p)$ is thus proportional to

$$\mathbb{P}\big(\delta|p| \leq \cos(\theta_k) \leq \delta(|p|+1)\big).$$

We start with the specific case of dimension $D = 2$, and we refer the reader to Fig. 6. In this case, θ is uniformly distributed on $[0, 2\pi]$. Therefore, for any $p \in \mathbb{N}$, $\mathbb{P}(\delta p \leq \cos(\theta) \leq \delta(p+1))$ is proportional to

$$\Pi_{\delta,p} = \int_{\arccos(\delta(p+1))}^{\arccos(\delta p)} \mathrm{d}\theta = \arccos(\delta p) - \arccos(\delta(p+1))$$

and $M_p = \#\{V_k \in S_p\}$ is driven by the binomial law with parameters $\mathscr{N}$ and $\alpha\Pi_{\delta,p}$ for a certain constant $\alpha > 0$. Hence, the analog of (2.7) is dominated, up to some constant, by

$$\mathscr{N}\Pi_{\delta,0} = \mathscr{N}\left(\frac{\pi}{2} - \arccos(\delta)\right) = \mathscr{N}\int_0^\delta \frac{\mathrm{d}x}{\sqrt{1-x^2}} \leq C\mathscr{N}\delta$$

as far as $0 < \delta \le \delta_0 < 1$. Similarly, the analog of (2.6) involves the sum

$$\sum_{p\ge 1} \frac{\mathscr{N}}{\delta^2 p^2} \Pi_{\delta,p}$$

that we split into

$$\mathrm{I} = \sum_{1\le p\le 1/2\delta} \ldots, \qquad \mathrm{II} = \sum_{1/2\delta<p\le 1/\delta} \ldots$$

For I, we can still use the fact that $x \mapsto \frac{1}{\sqrt{1-x^2}}$ is non increasing and bounded far away from $x = 1$ and we are led to the estimate

$$\begin{aligned} \mathrm{I} &= \sum_{1\le p\le 1/2\delta} \frac{\mathscr{N}}{\delta^2 p^2} \int_{\delta(p+1)}^{\delta p} \frac{\mathrm{d}x}{\sqrt{1-x^2}} \\ &\le \sum_{1\le p\le 1/2\delta} \frac{\mathscr{N}}{\delta^2 p^2} \frac{\delta}{\sqrt{1-\delta^2(p+1)^2}} \le C\frac{\mathscr{N}}{\delta}. \end{aligned}$$

For II, we use a summation by parts which yields

$$\begin{aligned} \mathrm{II} &= \sum_{1/2\delta<p\le 1/\delta} \frac{\mathscr{N}\arccos(\delta p)}{\delta^2} \Big(\frac{1}{(p-1)^2} - \frac{1}{p^2}\Big) \\ &\le \sum_{1/2\delta<p\le 1/\delta} \frac{\mathscr{N}\arccos(\delta p)}{\delta^2} \frac{2}{p(p-1)^2} \le \frac{4\delta}{\delta^2}\pi\mathscr{N} \sum_{p\ge 1} \frac{1}{p^2} \le C\frac{\mathscr{N}}{\delta}. \end{aligned}$$

Having these estimates at hand, we can repeat the same arguments as in the proof of Theorem 2.2.

For higher dimension, the situation is actually simpler since θ is now distributed on $[0, \pi/2]$ according to the law with density $\sin(\theta)^{D-2}\,\mathrm{d}\theta$. Thus (with the simple estimate $0 \le \sin(\theta)^{D-2} \le \sin(\theta)$) we obtain directly the analog of estimates (2.6) and (2.7). □

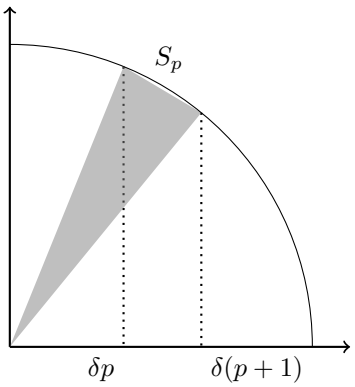

Figure 6: Velocities on the sphere $\mathbb{S}^1$, domain S_p.

The result can be extended to the L^p cases for $1 < p < \infty$ by using an interpolation argument as in [16, Theorem 2].

Corollary 2.4. *In Theorems 2.2 and 2.3, we assume that f and g belong to $L^p(\mathbb{R}^D \times \mathscr{V} \times \Omega, \mathrm{d}x\,\mathrm{d}\mu(v)\,\mathrm{d}\mathbb{P})$ for some $1 < p < \infty$, with $\mathscr{V}$ either $\mathbb{R}^D$ or $\mathbb{S}^{D-1}$. Then $\mathbb{E}\rho[\psi]$ lies in the Sobolev space $W^{s,p}(\mathbb{R}^D)$ with $0 < s < \min(1/p, 1-1/p) < 1$.*

Proof. We readily adapt the interpolation argument in [16]. Let $\mathscr{T}$ be the operator

$$\mathscr{T} : h \longmapsto \mathbb{E}\int f(x,v)\psi(v)\,\mathrm{d}\mu(v),$$

where

$$f(x,V_k) + V_k \cdot \nabla_x f(x,V_k) = h(x,V_k).$$

Clearly $\mathscr{T}$ maps continuously $L^r(\mathbb{R}^D \times \mathscr{V} \times \Omega, \, \mathrm{d}x \, \mathrm{d}\mu(v) \, \mathrm{d}\mathbb{P})$ into $L^r(\mathbb{R}^D)$, for any $1 < r < \infty$. Moreover, Theorems 2.2 and 2.3 tell us that $\mathscr{T}$ is a continuous operator from $L^2(\mathbb{R}^D \times \mathscr{V} \times \Omega, \, \mathrm{d}x \, \mathrm{d}\mu(v) \, \mathrm{d}\mathbb{P})$ to $H^{1/2}(\mathbb{R}^D)$. We conclude by interpreting the Sobolev space $W^{s,p}$ by interpolation, as being an intermediate space between $L^r = W^{0,r}$ and $H^{1/2} = W^{1/2,2}$ [4, Theorem 6.4.5, relation (7)], and L^p as being interpolated between L^r and L^2. □

We can equally extend the compactness statement to the L^1 framework, by following [17].

Corollary 2.5. *We consider a random set of velocities defined as inTheorem 2.2 or 2.3. Let* $\big(f_n\big)_{n\in\mathbb{N}}$ *and* $\big(g_n\big)_{n\in\mathbb{N}}$ *be two sequences of functions defined on* $\mathbb{R}^D \times \mathscr{V} \times \Omega$ *such that*

i) $\big\{f_n, \ n \in \mathbb{N}\big\}$ *is a relatively weakly compact set in* $L^1(\mathbb{R}^D \times \mathscr{V} \times \Omega, \, \mathrm{d}x \, \mathrm{d}\mu(v) \, \mathrm{d}\mathbb{P})$,

ii) $\big\{g_n, \ n \in \mathbb{N}\big\}$ *is bounded in* $L^1(\mathbb{R}^D \times \mathscr{V} \times \Omega, \, \mathrm{d}x \, \mathrm{d}\mu(v) \, \mathrm{d}\mathbb{P})$,

iii) we have $V_k \cdot \nabla_x f_n(x, V_k) = g_n(x, V_k)$.

Then $\mathbb{E}\rho_n[\psi](x) = \mathbb{E} \int f_n(x, v)\psi(v) \, \mathrm{d}\mu(v)$ *lies in a relatively compact set of* $L^1(B(0, R))$, *for any* $0 < R < \infty$ *(for the strong topology).*

Proof. The proof follows closely [17]; we sketch the arguments for the sake of completeness. For $\psi \in C_c^\infty(\mathscr{V})$, we denote by $\mathscr{A}$ the operator

$$\mathscr{A} : f \longmapsto \mathbb{E} \int f(x, v)\psi(v) \, \mathrm{d}\mu(v).$$

For $\lambda > 0$, we also introduce the operator

$$R_\lambda : h \longmapsto \int_0^\infty e^{-\lambda t} h(x - vt, v) \, \mathrm{d}t$$

which returns the solution $f = R_\lambda h$ of $(\lambda + v \cdot \nabla_x)f = h$. It is a continuous operator on $L^p(\mathbb{R}^D \times \mathscr{V}, \, \mathrm{d}x \, \mathrm{d}\mu(v))$ spaces and we have

$$\|R_\lambda h\|_{L^p} \leq \frac{\|h\|_{L^p}}{\lambda}. \tag{2.8}$$

Let us temporarily assume that the compactness statement holds for $\mathscr{A}R_\lambda g_n$, for any $\lambda > 0$, when i)-ii) is strengthened into

ii') $\quad \big\{g_n, \ n \in \mathbb{N}\big\}$ is a relatively weakly compact set in $L^1(\mathbb{R}^D \times \mathscr{V} \times \Omega, \, \mathrm{d}x \, \mathrm{d}\mu(v) \, \mathrm{d}\mathbb{P})$.

Therefore, writing $(\lambda + v \cdot \nabla_x)R_\lambda f_n = f_n$, we deduce from i) that $\big(\mathscr{A}R_\lambda f_n\big)_{n\in\mathbb{N}}$ is relatively compact in $L^1(B(0, R))$ for any $\lambda > 0$ and $0 < R < \infty$. Next, we write $f_n = \lambda R_\lambda f_n + R_\lambda(v \cdot \nabla_x f_n)$ so that, owing to (2.8), $\mathscr{A}f_n = \lambda\mathscr{A}R_\lambda f_n + \mathscr{A}R_\lambda(v \cdot \nabla_x f_n)$ appears as the sum of a sequence which is compact in $L^1(B(0, R))$ and a sequence the norm of which is dominated by $1/\lambda$, uniformly with respect to n. Consequently, $\big(\mathscr{A}f_n\big)_{b\in\mathbb{N}}$ is relatively compact in $L^1(B(0, R))$.

We are thus left with the task of justifying the gain of compactness for $\mathscr{A}R_\lambda g_n$ when i)-ii) is replaced by ii'), see [16, Proposition 3]. To this end, for $\lambda, M > 0$ we set $R_\lambda g_n = \gamma_n$ and we split

$$\gamma_n = \gamma_{n,M} + \gamma_n^M$$

where

$$\begin{aligned}(\lambda + V_k \cdot \nabla_x)\gamma_{n,M}(x, V_k) &= g_n(x, V_k)\mathbf{1}_{g_n(x,V_k)\leq M}, \\ (\lambda + V_k \cdot \nabla_x)\gamma_n^{M,\omega}(x, V_k) &= g_n(x, V_k)\mathbf{1}_{g_n(x,V_k)>M}.\end{aligned}$$

Since for any fixed $M > 0$, the set $\big\{g_n\mathbf{1}_{h_n\leq M}, \ n \in \mathbb{N}\big\}$ is bounded in $L^1 \cap L^\infty \subset L^2$, we can apply Theorems 2.2 or 2.3 which imply that $\big(\mathscr{A}\gamma_{n,M}\big)_{n\in\mathbb{N}}$ is compact in $L^1(B(0, R))$ for any finite R. We can conclude by showing that γ_n^M can be made arbitrarily small, in L^1 norm, uniformly with respect to $n \in \mathbb{N}$, for a suitable choice of $M > 0$. This is indeed the case because ii') implies

$$\lim_{M\to\infty} \left\{\sup_n \int |g_n|\mathbf{1}_{g_n>M} \, \mathrm{d}\mu(v) \, \mathrm{d}x \, \mathrm{d}\mathbb{P}(\omega)\right\} = 0.$$

by virtue of the Dunford-Pettis theorem, see [19, Sect. 7.3.2]. Going back to (2.8) finishes the proof. □

3 Application to the Rosseland Approximation

Let us go back to the asymptotic behavior of the solutions of (1.1). The problem (1.1) is completed with the initial condition

$$f_\varepsilon\Big|_{t=0} = f_\varepsilon^0.$$

It satisfies $f_\varepsilon^0 \geq 0$ and $f_\varepsilon^0 \in L^1(\mathbb{R}^D \times \mathscr{V})$, as it is physically relevant, f_ε being a particle density. For the set $(\mathscr{V}, \mathrm{d}\mu)$, in what follows we suppose at least that $\mathscr{V}$ is a bounded subset in $\mathbb{R}^D$ and

$$\int_{\mathscr{V}} \mathrm{d}\mu(v) = 1, \qquad \int_{\mathscr{V}} v \,\mathrm{d}\mu(v) = 0.$$

These assumptions are crucial for the analysis of the diffusion regime. Then, the connection to (1.2) can be established as follows.

Theorem 3.1. *We assume that (1.3) is fulfilled. Let σ be a function such that $\sigma(\rho) = \rho^\gamma \Sigma(\rho)$ with $|\gamma| < 1$ and $0 < \sigma_* \leq \Sigma(\rho) \leq \sigma^* < \infty$. Let $\left(f_\varepsilon^0\right)_{\varepsilon>0}$ satisfy*

$$\sup_{\varepsilon>0} \left(\int_{\mathbb{R}^d} \int_{\mathscr{V}} \left(1 + \varphi(x) + |\ln f_\varepsilon^0| f_\varepsilon^0\right) \mathrm{d}\mu(v)\,\mathrm{d}x + \|f_\varepsilon^0\|_{L^\infty(\mathbb{R}^d \times \mathscr{V})} \right) = M_0 < +\infty$$

for a certain weight function such that $\lim_{|x|\to+\infty} \varphi(x) = +\infty$. Then (up to a subsequence) the solution f_ε of (1.1) and ρ_ε converge to $\rho(t,x)$ in $L^p((0,T) \times \mathbb{R}^d \times \mathscr{V})$ and $L^p((0,T) \times \mathbb{R}^d)$ respectively, for any $1 \leq p < \infty$, $0 < T < \infty$ where ρ is a solution to (1.2) with the initial data $\rho\big|_{t=0}$ given by the weak limit in $L^p(\mathbb{R}^d)$ of $\int_{\mathscr{V}} f_\varepsilon^0 \,\mathrm{d}\mu(v)$ as $\varepsilon \to 0$.

For instance this statement holds with $\mathscr{V} = \mathbb{S}^{D-1}$ endowed with the Lebesgue measure. We refer the reader to [3] for a detailed proof, where the velocity averaging lemma is used to manage the passage to the limit in the nonlinearity. Assumption (1.3) can be replaced by

$$\text{for any } \xi \neq 0, \ \mathrm{meas}\Big(\Big\{v \in \mathscr{V} \cap B(0,R),\ v \cdot \xi \neq 0\Big\}\Big) > 0$$

which allows us to deal with certain discrete velocity models. Then, the asymptotic regime can be analyzed with a *compensated compactness* argument, that relies on the structure of the system satisfied by the zeroth and first moments of f_ϵ, as pointed out in [10, 20, 23], see also [24]. The question can be addressed of the relation between the diffusion equation that correspond to a discretization of the velocity set (discrete ordinate equation) and the diffusion equation that corresponds to the continuous model. For the simple collision operator in (1.1), velocity grids, which differ from the simplest uniform mesh, can be constructed that lead to the *exact* diffusion coefficient (namely $\frac{1}{\mathscr{N}} \sum_{k=1}^{\mathscr{N}} v_k \otimes v_k = \int_{\mathbb{S}^{D-1}} v \otimes v \,\mathrm{d}v = \frac{1}{D}\mathbb{I}$); we refer the reader to [7, 15, 21] for further discussion on this issue. However, for more general collision operators, it might happen that the equilibrium functions that make the collision operator vanish or the diffusion coefficient are not explicitly known, see [6, 10].

We wish to revisit this question by means of a Monte–Carlo approach: instead of the discrete ordinate viewpoint where a discrete velocity grid is adopted once for all, we deal with a random set of velocities and we wonder whether it can provide, in expectation, a consistent approximation of the diffusion regime. The consistency analysis we propose uses Theorem 2.2 or Theorem 2.3 to justify the following claim.

Theorem 3.2. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $V_1, ..., V_{\mathscr{N}}$ be i.i.d. random variables distributed according to the continuous uniform law on $\mathscr{V}$. Then, we obtain a set $\mathscr{V}_{\mathscr{N}}$ of $2\mathscr{N}$ velocities in $\mathscr{V}$ by setting $V_{\mathscr{N}+j} = -V_j$, for all $j \in \{1, ..., \mathscr{N}\}$. We denote the associated discrete measure on $\mathscr{V}$ by*

$$\mathrm{d}\mu_{\mathscr{N}}(v) = \frac{1}{2\mathscr{N}} \sum_{k=1}^{2\mathscr{N}} \delta(v = V_k).$$

Let $f_\varepsilon\big|_{t=0} = f_\varepsilon^0 \geq 0$ satisfy

$$\sup_{\varepsilon>0,\ \mathscr{N}\in\mathbb{N}} \left(\mathbb{E} \int_{\mathbb{R}^D} \int_{\mathscr{V}} (1 + \varphi(x) + |\ln f_\varepsilon^0|) f_\varepsilon^0 \,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x + \|f_\varepsilon^0\|_{L^\infty(\Omega \times \mathbb{R}^d \times \mathscr{V})} \right) = M_0 < +\infty. \tag{3.1}$$

Let f_ε be a solution of the following equation

$$\partial_t f_\varepsilon(t,x,V_j) + \frac{1}{\varepsilon} V_j \cdot \nabla_x f_\varepsilon(t,x,V_j) = \frac{1}{\varepsilon^2}\sigma(\rho_{\varepsilon,\mathscr{N}})\left[\rho_{\varepsilon,\mathscr{N}}(t,x) - f_\varepsilon(t,x,V_j)\right], \tag{3.2}$$

with $\rho_{\varepsilon,\mathscr{N}}(t,x) := \frac{1}{2\mathscr{N}}\sum_{i=1}^{2\mathscr{N}} f_\varepsilon(t,x,V_j)$. We suppose that $\rho \in [0,\infty) \mapsto \sigma(\rho)$ is a nonnegative function such that for any $0 < R < \infty$, there exists $\sigma_\star(R) > 0$ verifying $0 < 1/\sigma_\star(R) \leq \sigma(\rho) \leq \sigma_\star(R)$ and $|\sigma'(\rho)| \leq \sigma_\star(R)$ for any $0 \leq \rho \leq R$. Then $\mathbb{E}\rho_{\varepsilon,\mathscr{N}}$ converges to $\mathbb{E}\rho_{\mathscr{N}}$ in $L^2((0,T)\times\mathbb{R}^D)$ as ε goes to 0 with $0 < T < \infty$ where $\mathbb{E}\rho_{\mathscr{N}}$ is solution of

$$\begin{aligned}
&\partial_t \mathbb{E}\rho_{\mathscr{N}} + \operatorname{div}(\mathscr{J}_{\mathscr{N}}) = 0,\\
&\sigma(\mathbb{E}\rho_{\mathscr{N}})\mathscr{J}_{\mathscr{N}} = -\mathbb{E}A_{\mathscr{N}}\nabla_x \mathbb{E}\rho_{\mathscr{N}} + O\left(\frac{1}{\sqrt{\mathscr{N}}}\right),
\end{aligned}$$

with $A_{\mathscr{N}}$ the $D\times D$ matrix with random components defined by

$$A_{\mathscr{N}} := \frac{1}{2\mathscr{N}}\sum_{j=1}^{2\mathscr{N}} V_j \otimes V_j,$$

and $\mathbb{E}\rho_{\mathscr{N}}\big|_{t=0}$ is the weak limit of $\int \mathbb{E}f_\varepsilon^0 \,\mathrm{d}\mu(v)$.

Note that the construction of the set $\mathscr{V}_{\mathscr{N}}$ ensures that the null flux condition $\int v \,\mathrm{d}\mu_{\mathscr{N}}(v) = 0$ is fulfilled, but the elements of $\mathscr{V}_{\mathscr{N}}$ are not independent. Nevertheless, the stochastic averaging lemma still applies to this situation, with a straightforward adaptation of the proof. It is likely that the assumptions on σ can be substantially weakened, but this not our aim here to seek refinements in this direction. We will make precise in the proof in which sense the consistency error $O\left(\frac{1}{\sqrt{\mathscr{N}}}\right)$ should be understood.

3.1 Entropy estimates

In order to prove Theorem 3.2, the first step consists in establishing some a priori estimates, uniform with respect to the parameters ε and $\mathscr{N}$. We will then deduce the compactness needed to obtain the result. These estimates are quite classical; the proof that we sketch for the sake of completeness follows directly from [3, 20, 23].

Proposition 3.3. *Let f_ε^0 satisfy (3.1) with $\varphi(x) = (1+x^2)^\beta$, $0 < \beta < 1$. Let $0 < T < \infty$. There exists a constant $C(T)$ which only depends on T such that*

$$\begin{aligned}
\sup_{\varepsilon>0,\ \mathscr{N}\in\mathbb{N}} \Big\{ \sup_{0\leq t\leq T} \mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} (1+\varphi(x)+|\ln f_\varepsilon|) f_\varepsilon \,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \\
+\|f_\varepsilon\|_{L^\infty(\Omega\times(0,T)\times\mathbb{R}^D\times\mathscr{V})}\Big\} = C(T) < +\infty
\end{aligned} \tag{3.3}$$

and, furthermore,

$$\sup_{\varepsilon>0,\ \mathscr{N}\in\mathbb{N}} \mathbb{E}\int_0^T\int_{\mathbb{R}^D}\int_{\mathscr{V}} \frac{\sigma(\rho_{\varepsilon,\mathscr{N}})}{\varepsilon^2}(f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln\left(\frac{f_\varepsilon}{\rho_{\varepsilon,\mathscr{N}}}\right)\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x\,\mathrm{d}t \leq C(T). \tag{3.4}$$

Proof. As said above we crucially use the fact that

$$\int_{\mathscr{V}} \mathrm{d}\mu_{\mathscr{N}}(v) = 1, \qquad \int_{\mathscr{V}} v\,\mathrm{d}\mu_{\mathscr{N}}(v) = 0.$$

As a matter of fact, the collision operator is mass–conserving in the sense that

$$\int_{\mathscr{V}} \sigma(\rho)(f-\rho)\,\mathrm{d}\mu_{\mathscr{N}}(v) = 0.$$

Accordingly, integrating immediately leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon \,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x = 0. \tag{3.5}$$

More generally, let $G : [0,\infty) \to \mathbb{R}$ be a convex function. We get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} G(f_\varepsilon)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \\ = -\frac{1}{\varepsilon^2}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \sigma(\rho_{\varepsilon,\mathscr{N}})(\rho_{\varepsilon,\mathscr{N}} - f_\varepsilon)(G'(\rho_{\varepsilon,\mathscr{N}}) - G'(f_\varepsilon))\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \le 0.\end{aligned}$$

With $G(z) = z^p$, $p \ge 1$, it gives an estimate on the L^p norm of the solution. Similarly, with $G(z) = [z - \|f_\varepsilon^0\|_\infty]_+^2$, we conclude that

$$\|f_\varepsilon\|_{L^\infty(\Omega\times(0,T)\times\mathbb{R}^D\times\mathscr{V})} \le \|f_\varepsilon^0\|_\infty.$$

Finally, with $G(z) = z\ln(z)$ we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon \ln f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \\ = -\frac{1}{\varepsilon^2}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \sigma(\rho_{\varepsilon,\mathscr{N}})\,[\rho_{\varepsilon,\mathscr{N}} - f_\varepsilon]\ln\left(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}}\right)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \le 0.\end{aligned} \tag{3.6}$$

Let us focus on the following quantity obtained by multiplying (3.2) by φ and integrating

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi(x) f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x &= -\frac{1}{\varepsilon}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi(x) v\cdot\nabla_x f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \\ &= \frac{1}{\varepsilon}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon v\cdot\nabla_x\varphi(x)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \\ &= \mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} v\cdot\nabla_x\varphi(x)\frac{f_\varepsilon - \rho_{\varepsilon,\mathscr{N}}}{\varepsilon}\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x.\end{aligned}$$

Note that we have used $\int v\,\mathrm{d}_{\mathscr{N}}(v) = 0$. By Cauchy-Schwarz inequality, we know that

$$|\sqrt{b} - \sqrt{a}|^2 = \left|\int_a^b \frac{\mathrm{d}s}{2\sqrt{s}}\right|^2 \le \left|\int_a^b \frac{\mathrm{d}s}{4s}\right|\left|\int_a^b \mathrm{d}s\right| = \frac{1}{4}(b-a)\ln(b-a).$$

Thus, we get

$$\begin{aligned}\int_{\mathscr{V}} |f_\varepsilon - \rho_{\varepsilon,\mathscr{N}}|\,\mathrm{d}\mu_{\mathscr{N}}(v) &= \int_{\mathscr{V}} (\sqrt{f_\varepsilon} + \sqrt{\rho_{\varepsilon,\mathscr{N}}})\left|\sqrt{f_\varepsilon} - \sqrt{\rho_{\varepsilon,\mathscr{N}}}\right|\,\mathrm{d}\mu_{\mathscr{N}}(v) \\ &\le \left(\int_{\mathscr{V}} (\sqrt{f_\varepsilon} + \sqrt{\rho_{\varepsilon,\mathscr{N}}})^2\,\mathrm{d}\mu_{\mathscr{N}}(v)\right)^{1/2}\left(\int_{\mathscr{V}} (\sqrt{f_\varepsilon} - \sqrt{\rho_{\varepsilon,\mathscr{N}}})^2\,\mathrm{d}\mu_{\mathscr{N}}(v)\right)^{1/2} \\ &\le C\sqrt{\rho_{\varepsilon,\mathscr{N}}}\left(\int_{\mathscr{V}} (f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}})\,\mathrm{d}\mu_{\mathscr{N}}(v)\right)^{1/2}\end{aligned}$$

and we finally obtain the following bound

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \le \|v\|_{L^\infty(\Omega\times S)}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} |\nabla_x\varphi\frac{|f_\varepsilon - \rho_{\varepsilon,\mathscr{N}}|}{\varepsilon}\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \\ &\le C\ \mathbb{E}\int_{\mathbb{R}^D} |\nabla_x\varphi|\sqrt{\frac{\rho_{\varepsilon,\mathscr{N}}}{\sigma(\rho_{\varepsilon,\mathscr{N}})}}\left(\int_{\mathscr{V}} \frac{\sigma(\rho_{\varepsilon,\mathscr{N}})}{\varepsilon^2}(f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}})\,\mathrm{d}\mu_{\mathscr{N}}(v)\right)^{1/2}\mathrm{d}x \\ &\le C\mathbb{E}\left(\int_{\mathbb{R}^D} |\nabla_x\varphi|^2\frac{\rho_{\varepsilon,\mathscr{N}}}{\sigma(\rho_{\varepsilon,\mathscr{N}})}\,\mathrm{d}x\right)^{1/2} \\ &\qquad\times\left(\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \frac{\sigma(\rho_{\varepsilon,\mathscr{N}})}{\varepsilon^2}(f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}})\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x\right)^{1/2}.\end{aligned}$$

By assumption $1/\sigma(\rho_{\varepsilon,\mathscr{N}})$ is uniformly bounded. It follows that

$$\begin{aligned}\mathbb{E}\int_{\mathbb{R}^D} |\nabla_x\varphi|^2\frac{\rho_{\varepsilon,\mathscr{N}}}{\sigma(\rho_{\varepsilon,\mathscr{N}})}\,\mathrm{d}x &\le C\left(\mathbb{E}\int_{\mathbb{R}^D} |\nabla_x\varphi|^{2q}\,\mathrm{d}x\right)^{1/q}\left(\mathbb{E}\int_{\mathbb{R}^D} \rho_{\varepsilon,\mathscr{N}}^p\,\mathrm{d}x\right)^{1/p} \\ &\le C\left(\mathbb{E}\int_{\mathbb{R}^D} |\nabla_x\varphi|^{2q}\,\mathrm{d}x\right)^{1/q}\left(\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} |f_\varepsilon|^p\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x\right)^{1/p} \\ &\le C\end{aligned}$$

holds provided the Hölder conjugate q of $p \geq 1$ satisfies $\beta \leq 1/2 - D/4q$.
The Young inequality $ab \leq \frac{a^2}{4\theta} + \theta b^2$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi(x) f_\varepsilon(t,x,v)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x$$
$$\leq C + \frac{1}{2}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \frac{\sigma(\rho_{\varepsilon,\mathscr{N}})}{\varepsilon^2}(f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}})\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x.$$

Let us set

$$D_\varepsilon := \mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \frac{\sigma(\rho_{\varepsilon,\mathscr{N}})}{\varepsilon^2}(f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}})\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x \geq 0.$$

Coming back to (3.6), we get

$$\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon(t,x,v)\ln f_\varepsilon(t,x,v)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x$$
$$+\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi(x) f_\varepsilon(t,x,v)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x + \frac{1}{2}\int_0^t D_\varepsilon(s)\,\mathrm{d}s$$
$$\leq Ct + \mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon^{\omega,0}(x,v)\ln f_\varepsilon^{\omega,0}(x,v)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x + \mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi(x) f_\varepsilon^{\omega,0}(x,v)\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x.$$

Since $z|\ln z| = z\ln z - 2z\ln z\mathbf{1}_{\{0\leq z\leq 1\}}$, we have

$$\begin{aligned} 0 &\leq -\int_{0\leq f\leq 1} f\ln f\,\mathrm{d}y = -\int_{0\leq f\leq e^{-\varphi}} f\ln f\,\mathrm{d}y - \int_{e^{-\varphi}\leq f\leq 1} f\ln f\,\mathrm{d}y \\ &\leq \int \varphi f\,\mathrm{d}y + \int e^{-\varphi/2}\,\mathrm{d}y. \end{aligned}$$

Then, we are led to

$$\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon|\ln f_\varepsilon|\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x + \frac{1}{2}\int_0^t D_\varepsilon(s)\,\mathrm{d}s + \frac{1}{2}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x$$
$$= \mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon\ln f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x - 2\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon\ln f_\varepsilon\mathbf{1}_{\{0\leq f_\varepsilon\leq 1\}}\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x$$
$$+\frac{1}{2}\int_0^t D_\varepsilon(s)ds + \frac{1}{2}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x$$

$$\leq \mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} f_\varepsilon\ln f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x + 2\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \frac{\varphi}{4} f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x$$
$$+2\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} e^{-\varphi/8}\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x + \frac{1}{2}\int_0^t D_\varepsilon(s)ds + \frac{1}{2}\mathbb{E}\int_{\mathbb{R}^D}\int_{\mathscr{V}} \varphi f_\varepsilon\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x$$

$$\leq C(T)$$

which ends the proof. □

Moreover, we can deduce from above that f_ε behaves like its macroscopic part $\rho_{\varepsilon,\mathscr{N}}$ for small ε's.

Corollary 3.4. *We set* $g_{\varepsilon,\mathscr{N}} := \dfrac{f_\varepsilon - \rho_{\varepsilon,\mathscr{N}}}{\varepsilon}$. *Then, we have*

$$\sup_{\varepsilon>0,\ \mathscr{N}} \mathbb{E}\int_0^T\int_{\mathbb{R}^D}\left|\int_{\mathscr{V}} g_{\varepsilon,\mathscr{N}}\,\mathrm{d}\mu_{\mathscr{N}}(v)\right|^2 \mathrm{d}x\,\mathrm{d}t \leq C(T).$$

Proof. We write

$$\mathbb{E}\int_0^T\int_{\mathbb{R}^D}\left|\int_{\mathscr{V}} g_{\varepsilon,\mathscr{N}}\,\mathrm{d}\mu_{\mathscr{N}}(v)\right|^2 \mathrm{d}xdt = \mathbb{E}\int_0^T\int_{\mathbb{R}^D}\left(\int_{\mathscr{V}} \frac{|f_\varepsilon - \rho_{\varepsilon,\mathscr{N}}|}{\varepsilon}\,\mathrm{d}\mu_{\mathscr{N}}(v)\right)^2 \mathrm{d}x\,\mathrm{d}t$$
$$\leq C\mathbb{E}\int_0^T\int_{\mathbb{R}^D} \rho_{\varepsilon,\mathscr{N}}\int_{\mathscr{V}} (f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}})\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x\,\mathrm{d}t$$
$$\leq C\mathbb{E}\int_0^T\int_{\mathbb{R}^D} \frac{\rho_{\varepsilon,\mathscr{N}}}{\sigma(\rho_{\varepsilon,\mathscr{N}})}\int_{\mathscr{V}} \sigma(\rho_{\varepsilon,\mathscr{N}})(f_\varepsilon - \rho_{\varepsilon,\mathscr{N}})\ln(f_\varepsilon/\rho_{\varepsilon,\mathscr{N}})\,\mathrm{d}\mu_{\mathscr{N}}(v)\,\mathrm{d}x\,\mathrm{d}t.$$

Since by assumption on σ we know that $z \mapsto \frac{z}{\sigma(z)}$ is bounded on bounded sets and since $\rho_{\varepsilon,\mathscr{N}}$ is bounded in $L^\infty(\Omega\times(0,T)\times\mathbb{R}^D)$, we can conclude by using (3.4). □

3.2 Diffusive limit

We can now discuss how to pass to the limit $\varepsilon \to 0$.

Proof of Theorem 3.2. Applying Dunford-Pettis' theorem, see [19, Sect. 7.3.2] we deduce from Proposition 3.3 that, possibly at the price of extracting a subsequence,

$$f_\varepsilon \rightharpoonup f_{\mathcal{N}} \text{ weakly in } L^1(\Omega \times (0,T) \times \mathbb{R}^D \times \mathscr{V}_{\mathcal{N}}).$$

Consequently, we also have

$$\rho_{\varepsilon,\mathcal{N}} = \int_{\mathscr{V}} f_\varepsilon \, \mathrm{d}\mu_{\mathcal{N}}(v) \rightharpoonup \rho_{\mathcal{N}} = \int_{\mathscr{V}} f_{\mathcal{N}} \, \mathrm{d}\mu_{\mathcal{N}}(v) \text{ weakly in } L^1(\Omega \times (0,T) \times \mathbb{R}^D)$$

and

$$\mathbb{E}\rho_{\varepsilon,\mathcal{N}} \rightharpoonup \mathbb{E}\rho_{\mathcal{N}} \text{ weakly in } L^1((0,T) \times \mathbb{R}^D).$$

Next, we consider the equations satisfied by the moments of f_ε. To this end, let us set

$$J_{\varepsilon,\mathcal{N}}(t,x) := \frac{1}{2\mathcal{N}} \sum_{i=1}^{2\mathcal{N}} \frac{V_i}{\varepsilon} f_\varepsilon(t,x,V_i), \qquad \mathbb{P}_{\varepsilon,\mathcal{N}}(t,x) := \frac{1}{2\mathcal{N}} \sum_{i=1}^{2\mathcal{N}} V_i \otimes V_i f_\varepsilon(t,x,V_i).$$

Integrating (3.2) with respect to the velocity variable v yields

$$\partial_t \rho_{\varepsilon,\mathcal{N}} + \operatorname{div}(J_{\varepsilon,\mathcal{N}}) = 0. \tag{3.7}$$

Similarly, multiplying (3.2) by v and integrating leads to

$$\varepsilon^2 \partial_t J_{\varepsilon,\mathcal{N}} + \operatorname{div}(\mathbb{P}_{\varepsilon,\mathcal{N}}) = -\sigma(\rho_{\varepsilon,\mathcal{N}}) J_{\varepsilon,\mathcal{N}}. \tag{3.8}$$

Lemma 3.1. *The sequence $\left(J_{\varepsilon,\mathcal{N}}\right)_{\varepsilon>0}$ is bounded in $L^2(\Omega \times (0,T) \times \mathbb{R}^D)$ and we can write $\mathbb{P}_{\varepsilon,\mathcal{N}} = A_{\mathcal{N}} \rho_{\varepsilon,\mathcal{N}} + \varepsilon \mathbb{K}_{\varepsilon,\mathcal{N}}$ with $A_{\mathcal{N}} = \frac{1}{2\mathcal{N}} \sum_{j=1}^{2\mathcal{N}} V_j \otimes V_j$ and the components of $\left(\mathbb{K}_{\varepsilon,\mathcal{N}}\right)_{\varepsilon>0}$ are bounded in $L^2(\Omega \times (0,T) \times \mathbb{R}^D)$.*

Proof. The proof is based on the fact that $f_\varepsilon = \rho_{\varepsilon,\mathcal{N}} + \varepsilon g_{\varepsilon,\mathcal{N}}$. Since $\sum_{j=1}^{2\mathcal{N}} V_j = 0$, it allows us to write

$$J_{\varepsilon,\mathcal{N}} = \int v g_{\varepsilon,\mathcal{N}} \, \mathrm{d}\mu_{\mathcal{N}}(v),$$

and we deduce the bound on $J_{\varepsilon,\mathcal{N}}$ from Corollary 3.4 since $\|v\|_{L^\infty(\Omega \times S)} \leq C$. In addition, we have

$$\mathbb{P}_{\varepsilon,\mathcal{N}} = \int v \otimes v \, \mathrm{d}\mu_{\mathcal{N}}(v) \rho_{\varepsilon,\mathcal{N}} + \varepsilon \int v \otimes v g_{\varepsilon,\mathcal{N}} \, \mathrm{d}\mu_{\mathcal{N}}(v).$$

We set

$$\mathbb{K}_{\varepsilon,\mathcal{N}}(t,x) := \int v \otimes v g_{\varepsilon,\mathcal{N}}(t,x,v) \, \mathrm{d}\mu_{\mathcal{N}}(v).$$

We conclude by using the estimates in Corollary 3.4 again. □

Owing to Lemma 3.1, (3.8) can be recast as

$$\varepsilon\big(\varepsilon \partial_t J_{\varepsilon,\mathcal{N}} + \operatorname{div}(\mathbb{K}_{\varepsilon,\mathcal{N}})\big) + A_{\mathcal{N}} \nabla_x \rho_{\varepsilon,\mathcal{N}} = -\nu_{\varepsilon,\mathcal{N}}$$

with $\nu_{\varepsilon,\mathcal{N}} := \sigma(\rho_{\varepsilon,\mathcal{N}}) J_{\varepsilon,\mathcal{N}}$. Passing to the limit, up to subsequences, we are led to

$$\left\{ \begin{array}{rcl} \partial_t \rho_{\mathcal{N}} + \operatorname{div}(J_{\mathcal{N}}) & = & 0, \\ A_{\mathcal{N}} \nabla \rho_{\mathcal{N}} & = & -\nu_{\mathcal{N}} \end{array} \right. \tag{3.9}$$

where $\nu_{\mathcal{N}}$ is the weak limit as $\varepsilon \to 0$ of $\nu_{\varepsilon,\mathcal{N}}$, which is a bounded sequence in $L^2(\Omega \times (0,T) \times \mathbb{R}^D)$. It remains to establish a relation between $\nu_{\mathcal{N}}$, $\rho_{\mathcal{N}}$ and $J_{\mathcal{N}}$, or more precisely the expectation of these quantities. To this end, we are going to use the strong compactness of $\mathbb{E}\rho_{\varepsilon,\mathcal{N}}$ by using the averaging lemma. Indeed, we know that $\mathbb{E}\rho_{\varepsilon,\mathcal{N}}$ belongs to a bounded set in $L^2(0,T; H^{1/2}(\mathbb{R}^D))$; the proof follows exactly the same argument as for Theorem 2.2 taking the Fourier transform with respect to both the time and space variables t, x. However, because of the ε in front of the time derivative, we can not expect a gain of regularity with respect to the time variable. Then, we need to combine this estimate with another argument as follows:

(i) by using the Weil-Kolmogorov-Fréchet theorem, see [19, Th. 7.56], we deduce from the averaging lemma that

$$\lim_{|h|\to 0}\left(\sup_{\varepsilon}\int_0^T\int_{\mathbb{R}^D}|\mathbb{E}\rho_{\varepsilon,\mathscr{N}}(t,x+h)-\mathbb{E}\rho_{\varepsilon,\mathscr{N}}(t,x)|^2\,\mathrm{d}x\,\mathrm{d}t\right)=0,$$

(ii) Going back to (3.7), Lemma 3.1 tells us that $\partial_t\mathbb{E}\rho_{\varepsilon,\mathscr{N}}=-\mathrm{div}(\mathbb{E}J_{\varepsilon,\mathscr{N}})$ is bounded, uniformly with respect to ε, in $L^2(0,T;H^{-1}(\mathbb{R}^D))$.

Then, this is enough to deduce that $\mathbb{E}\rho_{\varepsilon,\mathscr{N}}$ strongly converges to $\mathbb{E}\rho_{\mathscr{N}}$ in $L^2((0,T)\times\mathbb{R}^D)$ (see e.g. [2, Appendix B] for a detailed proof).

Then, we rewrite

$$\begin{aligned}\mathbb{E}J_{\varepsilon,\mathscr{N}}&=\mathbb{E}\left(\frac{\nu_{\varepsilon,\mathscr{N}}}{\sigma(\rho_{\varepsilon,\mathscr{N}})}\right)=\frac{\mathbb{E}\nu_{\varepsilon,\mathscr{N}}}{\sigma(\mathbb{E}\rho_{\varepsilon,\mathscr{N}})}+\mathbb{E}r_{\varepsilon,\mathscr{N}},\\ r_{\varepsilon,\mathscr{N}}&=\left[\nu_{\varepsilon,\mathscr{N}}\left(\frac{1}{\sigma(\rho_{\varepsilon,\mathscr{N}})}-\frac{1}{\sigma(\mathbb{E}\rho_{\varepsilon,\mathscr{N}})}\right)\right].\end{aligned}\tag{3.10}$$

From the previous discussion, extracting further subsequences if necessary, we know that $\mathbb{E}\nu_{\varepsilon,\mathscr{N}}$ converges weakly to $\mathbb{E}\nu_N$ in $L^2((0,T)\times\mathbb{R}^D)$ while $\mathbb{E}\rho_{\varepsilon,\mathscr{N}}$ converges strongly in $L^2((0,T)\times\mathbb{R}^D)$ and a.e. to $\mathbb{E}\rho_{\mathscr{N}}$. Since σ is continuous and bounded from below, $1/\sigma(\mathbb{E}\rho_{\varepsilon,\mathscr{N}})$ converges to $1/\sigma(\mathbb{E}\rho_{\mathscr{N}})$ a.e. too, and it is bounded in $L^\infty((0,T)\times\mathbb{R}^D)$. We deduce that

$$\frac{\mathbb{E}\nu_{\varepsilon,\mathscr{N}}}{\sigma(\mathbb{E}\rho_{\varepsilon,\mathscr{N}})}\rightharpoonup\frac{\mathbb{E}\nu_{\mathscr{N}}}{\sigma(\mathbb{E}\rho_{\mathscr{N}})}\quad\text{weakly in } L^2((0,T)\times\mathbb{R}^D).$$

We are left with the task of proving that the last term in the right hand side of (3.10) tends to 0 as $\mathscr{N}\to\infty$, uniformly with respect to ε. The Cauchy–Schwarz inequality yields

$$\begin{aligned}\left|\mathbb{E}r_{\varepsilon,\mathscr{N}}\right|&\le\left(\mathbb{E}\left[(\nu_{\varepsilon,\mathscr{N}})^2\right]\right)^{1/2}\left(\mathbb{E}\left[\left(\frac{1}{\sigma(\rho_{\varepsilon,\mathscr{N}})}-\frac{1}{\sigma(\mathbb{E}\rho_{\varepsilon,\mathscr{N}})}\right)^2\right]\right)^{1/2}\\ &\le\left(\mathbb{E}\left[(\nu_{\varepsilon,\mathscr{N}})^2\right]\right)^{1/2}\left(\mathbb{E}\left[\left(\int_{\mathbb{E}\rho_{\varepsilon,\mathscr{N}}}^{\rho_{\varepsilon,\mathscr{N}}}\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{1}{\sigma(z)}\right]\mathrm{d}z\right)^2\right]\right)^{1/2}\\ &\le\left(\mathbb{E}\left[(\nu_{\varepsilon,\mathscr{N}})^2\right]\right)^{1/2}\left(\mathbb{E}\left[(\rho_{\varepsilon,\mathscr{N}}-\mathbb{E}\rho_{\varepsilon,\mathscr{N}})^2\right]\right)^{1/2}\\ &\le\left(\mathbb{E}\left[(\nu_{\varepsilon,\mathscr{N}})^2\right]\right)^{1/2}\left(\mathbb{E}\left[\left(\frac{1}{2\mathscr{N}}\sum_{i=1}^{2\mathscr{N}}f_\varepsilon(V_i)-\mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right)^2\right]\right)^{1/2}.\end{aligned}\tag{3.11}$$

We remind the reader that the $2\mathscr{N}$ velocities are constructed by symmetry from $V_1,...,V_{\mathscr{N}}$ which are i.i.d. velocities in $[-0.5,0.5]^D$ and we write

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{2\mathscr{N}}\sum_{i=1}^{2\mathscr{N}}f_\varepsilon(V_i)-\mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right]^2\\ &=\mathbb{E}\left[\frac{1}{4\mathscr{N}^2}\sum_{i,j=1}^{\mathscr{N}}\left\{\left(f_\varepsilon(V_i)+f_\varepsilon(-V_i)-2\mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right)\left(f_\varepsilon(V_j)+f_\varepsilon(-V_j)-2\mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right)\right\}\right].\end{aligned}\tag{3.12}$$

When $i\neq j$, V_i and V_j are independent, which implies

$$\begin{aligned}&\mathbb{E}\left[(f_\varepsilon(V_i)+f_\varepsilon(-V_i)-2\mathbb{E}\rho_{\varepsilon,\mathscr{N}})(f_\varepsilon(V_j)+f_\varepsilon(-V_j)-2\mathbb{E}\rho_{\varepsilon,\mathscr{N}})\right]\\ &\qquad=\mathbb{E}\left[f_\varepsilon(V_i)+f_\varepsilon(-V_i)-2\mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right]\mathbb{E}\left[f_\varepsilon(V_j)+f_\varepsilon(-V_j)-2\mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right].\end{aligned}$$

Now, we use the fact that the V_i's are identically distributed so that

$$\begin{aligned}2\mathbb{E}\rho_{\varepsilon,\mathscr{N}}&=2\mathbb{E}\left(\frac{1}{2\mathscr{N}}\sum_{k=1}^{2\mathscr{N}}f_\varepsilon(V_k)\right)=\mathbb{E}\left(\frac{1}{\mathscr{N}}\sum_{k=1}^{\mathscr{N}}\left(f_\varepsilon(V_k)+f_\varepsilon(-V_k)\right)\right)\\ &=\frac{1}{\mathscr{N}}\sum_{k=1}^{\mathscr{N}}\left(\mathbb{E}f_\varepsilon(V_k)+\mathbb{E}f_\varepsilon(-V_k)\right)=\mathbb{E}f_\varepsilon(V_j)+\mathbb{E}f_\varepsilon(-V_j)\end{aligned}$$

for any $j \in \{1, ..., \mathscr{N}\}$. It follows that

$$\mathbb{E}\left[(f_\varepsilon(V_i) + f_\varepsilon(-V_i) - 2\mathbb{E}\rho_{\varepsilon,\mathscr{N}})\,(f_\varepsilon(V_j) + f_\varepsilon(-V_j) - 2\mathbb{E}\rho_{\varepsilon,\mathscr{N}})\right] = 0 \quad \text{when } i \neq j.$$

Going back to (3.12), we obtain

$$\mathbb{E}\left[\frac{1}{2\mathscr{N}}\sum_{i=1}^{2\mathscr{N}} f_\varepsilon(V_i) - \mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right]^2 = \mathbb{E}\left[\frac{1}{4\mathscr{N}^2}\sum_{i=1}^{\mathscr{N}}\Big(f_\varepsilon(V_i) + f_\varepsilon(-V_i) - 2\mathbb{E}\rho_{\varepsilon,\mathscr{N}}\Big)^2\right].$$

Since f_ε and $\rho_{\varepsilon,\mathscr{N}}$ are uniformly bounded, we conclude that the estimate

$$\mathbb{E}\left[\frac{1}{2\mathscr{N}}\sum_{i=1}^{2\mathscr{N}} f_\varepsilon(V_i) - \mathbb{E}\rho_{\varepsilon,\mathscr{N}}\right]^2 \leq \frac{C}{\mathscr{N}}$$

holds. Inserting this information in (3.11), we arrive at

$$\int_0^T \int_{\mathbb{R}^D} \left|\mathbb{E}r_{\varepsilon,\mathscr{N}}\right|^2 \mathrm{d}x\,\mathrm{d}t \leq \frac{C}{\mathscr{N}}\,\mathbb{E}\int_0^T \int_{\mathbb{R}^D} \nu_{\varepsilon,\mathscr{N}}^2\,\mathrm{d}x\,\mathrm{d}t,$$

which is thus of order $O(1/\mathscr{N})$, uniformly with respect to ε.

Therefore, we can let ε run to 0 in (3.10) and, for a suitable subsequence, we are led to

$$\mathbb{E}J_{\varepsilon,\mathscr{N}} \rightharpoonup \mathbb{E}J_{\mathscr{N}} = \frac{\mathbb{E}\nu_{\mathscr{N}}}{\sigma(\mathbb{E}\rho_{\mathscr{N}})} + r_{\mathscr{N}} \quad \text{weakly in } L^2((0,T)\times\mathbb{R}^D) \text{ with } \|r_{\mathscr{N}}\|_{L^2((0,T)\times\mathbb{R}^D)} \leq \frac{C}{\sqrt{\mathscr{N}}}.$$

Finally, we take the expectation in (3.9) and we get

$$\mathbb{E}\left(A_{\mathscr{N}}\nabla_x\rho_{\mathscr{N}}\right) = -\mathbb{E}\nu_{\mathscr{N}} = -\sigma(\mathbb{E}\rho_{\mathscr{N}})\mathbb{E}J_{\mathscr{N}} + \sigma(\mathbb{E}\rho_{\mathscr{N}})r_{\mathscr{N}}.$$

Note that the last term is still of order $O(1/\sqrt{\mathscr{N}})$ in the $L^2((0,T)\times\mathbb{R}^D)$ norm. By a reasoning similar as above, we check that, for any $i, j \in \{1, ..., D\}$,

$$\sqrt{\mathbb{E}\left[\Big([A_{\mathscr{N}}]_{ij} - \mathbb{E}[A_{\mathscr{N}}]_{ij}\Big)^2\right]} = O\left(\frac{1}{\sqrt{\mathscr{N}}}\right)$$

(this is the standard result about Monte–Carlo integration). It implies that we can find a constant $C > 0$, which only depends on the dimension D, such that for any $\xi \in \mathbb{R}^D$,

$$\mathbb{E}\left[\Big|(A_{\mathscr{N}}\xi - \mathbb{E}[A_{\mathscr{N}}\xi]\Big|^2\right] \leq \frac{C|\xi|^2}{\mathscr{N}}.$$

Then we get

$$\mathbb{E}\left(A_{\mathscr{N}}\nabla_x\rho_{\mathscr{N}}\right) = \mathbb{E}A_{\mathscr{N}}\ \nabla_x\mathbb{E}\rho_{\mathscr{N}} + s_{\mathscr{N}},$$

$$s_{\mathscr{N}} = \mathbb{E}\left[(A_{\mathscr{N}} - \mathbb{E}A_{\mathscr{N}})\nabla_x\rho_{\mathscr{N}}\right].$$

The remainder term should be analyzed in a weak sense, due to a lack of a priori regularity of $\nabla_x\rho_{\mathscr{N}}$ (we only know that the product $A_{\mathscr{N}}\nabla_x\rho_{\mathscr{N}}$ lies in L^2, but the invertibility of $A_{\mathscr{N}}$ is not guaranteed). We have, for any $\varphi \in C_c^\infty((0,T)\times\mathbb{R}^D)$,

$$\begin{aligned} |\langle\mathbb{E}s_{\mathscr{N}}|\varphi\rangle| &= \left|-\mathbb{E}\int_0^T\int_{\mathbb{R}^D}\rho_{\mathscr{N}}\ (A_{\mathscr{N}} - \mathbb{E}A_{\mathscr{N}})\nabla_x\varphi\,\mathrm{d}x\,\mathrm{d}t\right| \\ &\leq \left(\mathbb{E}\int_0^T\int_{\mathbb{R}^D}\rho_{\mathscr{N}}^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\left(\int_0^T\int_{\mathbb{R}^D}|\nabla_x\varphi|^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\frac{C}{\sqrt{\mathscr{N}}}. \end{aligned}$$

Owing to the estimates (3.3) in Proposition 3.3, it means that $s_{\mathscr{N}}$ is therefore of order $O(1/\sqrt{\mathscr{N}})$ in the $L^2(0,T;H^{-1}(\mathbb{R}^D))$−norm. □

Remark 3.1. *The random matrix $A_{\mathscr{N}}$ might be singular. However $\mathbb{E}A_{\mathscr{N}}$ is invertible. Indeed for any $\xi \neq 0$, we have $\mathbb{E}A_{\mathscr{N}}\xi\cdot\xi = \frac{1}{2\mathscr{N}}\sum_{j=1}^{2\mathscr{N}}\mathbb{E}\big[|V_j\cdot\xi|^2\big] \geq 0$. This quantity is actually positive since $\mathbb{P}(v\cdot\xi = 0) = 0$ for the continuous laws we are dealing with.*

References

[1] V Agoshkov. Spaces of functions with differential-difference characteristics and the smoothness of solutions of the transport equation. *Dokl. Akad. Nauk SSSR*, 276(6):1289–1293, 1984.

[2] A. Alonso, T. Goudon, and A. Vavasseur. Damping of particles interacting with a vibrating medium. Technical report, Inria, 2016.

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

[4] J. Bergh and J. Löfström. *Interpolation spaces*, volume 223 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1976.

[5] F. Berthelin and S. Junca. Averaging lemmas with a force term in the transport equation. *J. Math. Pures Appl.*, 93(2):113–131, 2019.

[6] V. Bonnaillie-Noël, J. A. Carillo, T. Goudon, and G. A. Pavliotis. Efficient numerical calculation of drift and diffusion coefficients in the diffusion approximation of kinetic equations. *IMA J. Numer. Anal.*, 2016. To appear.

[7] C. Buet, S. Cordier, B. Lucquin-Desreux, and S. Mancini. Diffusion limit of the Lorentz model: Asymptotic preserving schemes. *ESAIM: M2AN*, 36(4):631–655, 2002.

[8] A. Debussche, S. De Moor, and J. Vovelle. Diffusion limit for the radiative transfer equation perturbed by a Wiener process. *Kinetic and Related Models*, 8(3):467–492, 2015.

[9] A. Debussche, S. De Moor, and J. Vovelle. Diffusion limit for the radiative transfer equation perturbed by a Markov process. *As. Anal.*, 2016. To appear.

[10] P. Degond, T. Goudon, and F. Poupaud. Diffusion limit for nonhomogeneous and non-micro-reversible processes. *Indiana Univ. Math. J.*, 49(3):1175–1198, 2000.

[11] R. J. DiPerna and P.-L. Lions. Global weak solutions of Vlasov–Maxwell systems. *Comm. Pure Appl. Math.*, 42(6):729–757, 1989.

[12] R. J. DiPerna and P.-L. Lions. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Ann. of Math. (2)*, 130(2):321–366, 1989.

[13] R. J. DiPerna, P.-L. Lions, and Y. Meyer. L^p regularity of velocity averages. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 8(3-4):271–287, 1991.

[14] F. Golse. From kinetic to macroscopic models. In B. Perthame and L. Desvillettes, editors, *Kinetic equations and asymptotic theory*, volume 4 of *Series in Appl. Math*, pages 41–121. Gauthier-Villars, 2000.

[15] F. Golse, S. Jin, and D. Levermore. The convergence of numerical transfer schemes in diffusive regimes I: Discrete-ordinate method. *SIAM J. Numer. Anal.*, 36(5):1333–1369, 1999.

[16] F. Golse, P.-L. Lions, B. Perthame, and R. Sentis. Regularity of the moments of the solution of a transport equation. *J. Funct. Anal.*, 76:110–125, 1988.

[17] F. Golse and L. Saint-Raymond. Velocity averaging in L^1 for the transport equation. *C. R. Math. Acad. Sci. Paris*, 334(7):557–562, 2002.

[18] F. Golse and L. Saint-Raymond. The Navier-Stokes limit of the Boltzmann equation for bounded collision kernels. *Invent. Math.*, 155(1):81–161, 2004.

[19] T. Goudon. *Intégration. Intégrale de Lebesgue et introduction à l'analyse fonctionnelle.* Références Sciences. Ellipses, 2011.

[20] T. Goudon and F. Poupaud. Approximation by homogenization and diffusion of kinetic equations. *Comm. Partial Differential Equations*, 26(3-4):537–569, 2001.

[21] S. Jin and D. Levermore. The discrete-ordinate method in diffusive regimes. *Tranp. Th. Stat. Phys.*, 20(5-6):413–439, 1991.

[22] P.-L. Lions, B. Perthame, and P. E. Souganidis. Stochastic averaging lemmas for kinetic equations. In *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Année 2011–2012*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XXVI, 17. École Polytech., Palaiseau, 2013.

[23] P.-L. Lions and G. Toscani. Diffuse limit for finite velocity Boltzmann kinetic models. *Rev. Mat. Ib.*, 13:473–513, 1997.

[24] P. Marcati and A. Milani. The one-dimensional Darcy's law as the limit of a compressible Euler flow. *J. Differential Eq.*, 84(1):129–147, 1990.

[25] S. Mischler. Convergence of discrete-velocity schemes for the Boltzmann equation. *Arch. Rational Mech. Anal.*, 140:53–77, 1997.

[26] B. Perthame and P. E. Souganidis. A limiting case for velocity averaging. *Ann. Sci. École Norm. Sup. (4)*, 31(4):591–598, 1998.

[27] L. Saint Raymond. *Hydrodynamic limits of the Boltzmann equation*, volume 1971 of *Lect. Notes in Math.* Springer, 2009.

[28] E. Tadmor and T. Tao. Velocity averaging, kinetic formulations, and regularizing effects in quasi-linear PDEs. *Comm. Pure Appl. Math.*, 60(10):1488–1521, 2007.

[29] C. Villani. Limites hydrodynamiques de l'équation de Boltzmann (d'après C. Bardos, F. Golse, C. D. Levermore, P.-L. Lions, N. Masmoudi, L. Saint-Raymond). *Astérisque*, (282):Exp. No. 893, ix, 365–405, 2002. Séminaire Bourbaki, Vol. 2000/2001.