DIFFUSION LIMIT FOR THE RADIATIVE TRANSFER EQUATION PERTURBED BY A WIENER PROCESS

ARNAUD DEBUSSCHE
IRMAR, ENS Rennes, CNRS, UEB
av. Robert Schuman, F-35170 Bruz, France

SYLVAIN DE MOOR AND JULIEN VOVELLE
IRMAR, ENS Rennes
av. Robert Schuman, F-35170 Bruz, France

and
Université de Lyon, CNRS, Université Lyon 1, Institut Camille Jordan
43 boulevard du 11 novembre 1918, F-69622 Villeurbanne Cedex, France

(Communicated by Benoit Perthame)

Abstract. The aim of this paper is the rigorous derivation of a stochastic non-linear diffusion equation from a radiative transfer equation perturbed with a random noise. The proof of the convergence relies on a formal Hilbert expansion and the estimation of the remainder. The Hilbert expansion has to be done up to order 3 to overcome some difficulties caused by the random noise.

1. Introduction. In this paper, we are interested in the following non-linear equation

\[
\begin{align*}
\frac{df^\varepsilon}{dt} + \frac{1}{\varepsilon}a(v) \cdot \nabla_x f^\varepsilon &= \frac{1}{\varepsilon^2} \sigma(\overline{f})L(f^\varepsilon)dt + f^\varepsilon \circ QdW_t, \\
f^\varepsilon(0) &= \rho_{\text{in}}, \quad x \in \mathbb{T}^N, \quad v \in V,
\end{align*}
\]

(1)

where $V$ is an $N$-dimensional torus, $a : V \to \mathbb{R}^N$ and $\sigma : \mathbb{R} \to \mathbb{R}$. The notation $\overline{f}$ stands for the average over the velocity space $V$ of the function $f$, that is

$$\overline{f} = \int_V f \, dv.$$

The operator $L$ is a linear operator of relaxation which acts on the velocity variable $v \in V$ only. It is given by

$$L(f) := \overline{f} - f.$$ 

The noise term described thanks to $W$ is a cylindrical Wiener process on the Hilbert space $L^2(\mathbb{T}^N)$ and a covariance operator $Q$. This latter is a linear self-adjoint operator on $L^2(\mathbb{T}^N)$. This is a usual way to define a noise which is white in time and correlated in space. The spatial correlation can be written explicitly in terms of $Q$. The precise description of the problem setting will be given in the next section. In this paper, we investigate the behavior in the limit $\varepsilon \to 0$ of the solution $f^\varepsilon$ of (1).

2010 Mathematics Subject Classification. Primary: 35R60, 60H15; Secondary: 35B25, 85A25.

Key words and phrases. Kinetic equations, diffusion limit, stochastic partial differential equations, Hilbert expansion, radiative transfer.
Concerning the physical background in the deterministic case \((Q \equiv 0)\), the equation (1) describes the interaction between a surrounding continuous medium and a flux of photons radiating through it in the absence of hydrodynamical motion. The unknown \(f^\varepsilon(t, x, v)\) then stands for a distribution function of photons having position \(x\) and velocity \(v\) at time \(t\). The function \(\sigma\) is the opacity of the matter. When the surrounding medium becomes very large compared to the mean free paths \(\varepsilon\) of photons, the solution \(f^\varepsilon\) to (1) with \(Q \equiv 0\) is known to behave like \(\rho\) where \(\rho\) is the solution of the Rosseland equation

\[
\partial_t \rho - \text{div}_x (\sigma(\rho)^{-1} K \nabla_x \rho) = 0, \quad (t, x) \in [0, T] \times \mathbb{T}^N.
\]

with \(K := \int_v a(v) \otimes a(v) \, dv\). This is called the Rosseland approximation. In the deterministic case, the Rosseland approximation has been widely studied. In the paper of Bardos, Golse and Perthame [1], the Rosseland approximation for a slightly more general equation of radiative transfer type than (1) is derived. Using the so-called Hilbert’s expansion method, Bardos, Golse and Perthame prove a strong convergence of the solution to the radiative transfer equation to the solution to the Rosseland equation. In [2], the stationary and evolution Rosseland approximation are proved in a weaker sense with weakened hypothesis on the various parameters of the radiative transfer equation, in particular on the opacity function \(\sigma\) that is possibly not bounded from above.

In this paper, we investigate such an approximation where we have perturbed the deterministic equation by a spatially smooth multiplicative random noise of the form \(f^\varepsilon \circ QdW\). This represents a random creation or absorption of photons. Note that the noise is independent of the velocity variable. We also consider the limit \(\varepsilon \to 0\) and obtain a stochastic PDE at the limit. Our proof also relies on the Hilbert expansion method: we expand the solution \(f^\varepsilon\) to (1) as \(f^\varepsilon = \rho + \varepsilon f_1 + \varepsilon^2 f_2 + \varepsilon^3 f_3 + r^\varepsilon\) where \(\rho\) is the solution to the limit problem, \(f_1, f_2, f_3\) are three correctors to be defined appropriately and where \(r^\varepsilon\) denotes the remainder of the expansion. First, we prove that the correctors \((f_i)_{1 \leq i \leq 3}\) behave correctly in the space \(X := L^\infty(0, T; L^1(\Omega; L^1_{x,v}))\) so that \(\varepsilon f_1 + \varepsilon^2 f_2 + \varepsilon^3 f_3 = O(\varepsilon)\) in \(X\). This step requires some regularity on the limit solution \(\rho\) and we use the regularity result in [11]. Then, to conclude the proof, we estimate the remainder by mean of an \(\text{Itô}\) formula to show that \(r^\varepsilon\) is of order \(\varepsilon\) in \(X\). Note that an Hilbert expansion up to order 2 is usually sufficient in many well-known deterministic cases; here we need to push the expansion up to order 3 to overcome some difficulties caused by the noise term.

In another article ([12]), we consider the same problem but in the spirit of [10]. The random term there is not white in time but is of the form \(\frac{1}{2} m(x, \frac{x}{\varepsilon^2})\) for a smooth stationary Markov process \(m\). The limit equation is the same, as expected since this random process converges to \(QdW\), where \(Q\) is a covariance operator which can be expressed in terms of the driving process \(m\). In [10], the case of a constant opacity, and thus of a linear equation, is considered. In this work, the convergence in law of the solutions to a limit stochastic fluid equation is obtained by a generalization of the perturbed test-functions method.

In this present work, we consider a non-linear operator \(\sigma(\widetilde{f})Lf\), which can be seen as a simple non-linear perturbation of the classical linear relaxation operator \(L\) considered in [10]. Nevertheless, we consider that the noise is already in its limit form \(QdW\). In particular, we point out that the fact that the noise is already in an \(\text{Itô}\) form permits the application of the \(\text{Itô}\) formula and of the Hilbert expansion method. As a consequence, we are able to prove in this paper a stronger result of convergence.
of $f$ to $\rho$, namely a strong convergence in the space $X := L^\infty(0,T;L^1(\Omega;L^1_x,\rho))$ with rate $\varepsilon$.

We point out that, in the sequel, when proving existence and uniqueness for the problem (1), we use a stochastic averaging lemma which can be interesting by itself. It provides a gain in regularity by averaging over the velocity space for solutions to kinetic stochastic equations, see Lemma 4.2. The proof of this lemma is detailed in Appendix B; it is mainly based on an adaptation to a stochastic setting of the paper of Bouchut and Desvillettes [4]. Note that it differs from the stochastic averaging lemma of Lions, Perthame, Souganidis [16] where the noise acts on the values of the velocity $a(v)$.

The paper is organized as follows. In Section 2, we introduce the setting and the notations and give the main result to be proved, Theorem 2.2. In Section 3, we derive formally the limit equation. Finally, in Section 4, we provide the proof of the main result, which is divided in three main steps. First, we study the existence, uniqueness and regularity of the solutions to the radiative transfer equation (1) and to the stochastic Rosseland problem. Then we define and study the correctors of the Hilbert expansion. Finally, we estimate the remainder to conclude the proof.

2. Preliminaries and main result.

2.1. Notations and hypothesis. Let us now introduce the precise setting of equation (1). We work on a finite-time interval $[0,T]$, $T > 0$, and consider periodic boundary conditions for the space variable: $x \in \mathbb{T}^N$ where $\mathbb{T}^N$ is the $N$-dimensional torus. Regarding the velocity space $V$, we also consider periodic boundary conditions, that is $V = \mathbb{T}^N$, but we keep the notation $V$ to distinguish the velocity space from the space one.

For $p \in [1,\infty]$, the Lebesgue spaces $L^p(\mathbb{T}^N \times V)$ will be denoted by $L^p_{x,v}$ for short. The associated norm will be written $\|\cdot\|_{L^p_{x,v}}$. Similarly, we define the Lebesgue spaces $L^p_x$, $L^p_v$ and, if $k \in \mathbb{Z}$, the Sobolev spaces $W^{k,p}_{x,v}$ and $W^{k,p}_v$ or $H^k_{x,v}$ and $H^k_v$ when $p = 2$. The scalar product of $L^2_{x,v}$ will be denoted by $(\cdot,\cdot)$. We finally introduce, for $k \in \mathbb{N}$, the space $C^{0,k}([0,T] \times \mathbb{T}^N)$ constituted by the functions of the variables $(t,x) \in [0,T] \times \mathbb{T}^N$ which are continuous in time and $k$-times continuously differentiable in space.

Concerning the velocity mapping $a : V \to V$, we shall assume that it is $C^1$. We will use a stochastic version of averaging lemmas to prove the existence of the solution $f^\varepsilon$ to (1). To do so, we need to assume the following standard non-degeneracy condition (indicating that $v \mapsto a(v)$ is not stationary):

$$\forall \varepsilon > 0, \forall (\xi,\sigma) \in S^{N-1} \times \mathbb{R}, \text{ Leb} \left( \{ v \in V, |a(v) \cdot \xi + \sigma| < \varepsilon \} \right) \leq \varepsilon^\alpha, \quad (4)$$

for some $\alpha \in (0,1]$ and where Leb denotes the normalized Lebesgue measure on $V = \mathbb{T}^N$. Furthermore, we will suppose that the following null flux hypothesis holds

$$\int_V a(v) \, dv = 0. \quad (5)$$

Actually, Equation (1) is issued from a parabolic rescaling $(t,x) \mapsto (\varepsilon^{-2}t,\varepsilon^{-1}t)$ which makes sense only if (5) is satisfied: in the case where the average velocity would be non trivial, a hyperbolic scaling $(t,x) \mapsto (\varepsilon^{-1}t,\varepsilon^{-1}t)$ would have to be considered. The following diffusion matrix

$$\mathcal{K} := \int_V a(v) \otimes a(v) \, dv \quad (6)$$
appears in the Rosseland Equation (3). Under the non-degeneracy hypothesis (4), $\mathbf{K}$ is definite positive. Indeed, (4) implies in particular that, for all $\xi \in S^{N-1}$, $|a(v) \cdot \xi| > 0$ a.e. in $V$. In particular

$$\mathbf{K} \xi \cdot \xi = \int_V |a(v) \cdot \xi|^2 dv$$

cannot vanish.

Regarding the opacity function $\sigma : \mathbb{R} \to \mathbb{R}$, we assume that

(H1) There exist two positive constants $\sigma_*, \sigma^* > 0$ such that for almost all $x \in \mathbb{R}$, we have

$$\sigma_* \leq \sigma(x) \leq \sigma^*;$$

(H2) the function $\sigma$ is $C^3$ (i.e. of class $C^3$ with bounded derivatives up to order three). In particular $\sigma$ is Lipschitz continuous;

(H3) the mappings $x \mapsto \sigma(x)$ and $x \mapsto \sigma(x)x$ are respectively non-increasing and non-decreasing.

Finally, the initial condition $\rho_{in}$ is supposed to be a smooth non-negative function which does not depend on the variable $v \in V$. More precisely, it is assumed to be in $C^{3+\eta}(\mathbb{T}^N)$ for some $\eta > 0$.

2.2. The random noise. Regarding the stochastic term, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a stochastic basis with a complete, right-continuous filtration. The random noise $dW_t$ is a cylindrical Wiener process on the Hilbert space $L^2(\mathbb{T}^N)$. We can define it by setting

$$dW_t = \sum_{k \geq 0} e_k \, d\beta_k(t), \quad (7)$$

where the $(\beta_k)_{k \geq 0}$ are independent Brownian motions on the real line and $(e_k)_{k \geq 0}$ a complete orthonormal system in the Hilbert space $L^2(\mathbb{T}^N)$. The covariance operator $Q$ is a linear self-adjoint operator on $L^2(\mathbb{T}^N)$. We assume the following regularity property

$$\sum_{k \geq 0} \|Qe_k\|^2_{L^\infty} < \infty. \quad (8)$$

In particular, we define

$$\kappa_{0,\infty} := \sum_{k \geq 0} \|Qe_k\|^2_{L^\infty} < \infty, \quad \kappa_{1,\infty} := \sum_{k \geq 0, 1 \leq i \leq N} \|\partial x_i Qe_k\|^2_{L^\infty} < \infty. \quad (9)$$

As a consequence, we can introduce

$$G := \frac{1}{2} \sum_{k \geq 0} (Qe_k)^2,$$

which will be useful when switching Stratonovich integrals into Ito form. Precisely, we point out that for Equation (1) we can write $f^\varepsilon \circ QdW_t = f^\varepsilon QdW_t + Gf^\varepsilon dt$ where

$$QdW_t = \sum_{k \geq 0} Qe_k \, d\beta_k(t).$$

In the sequel, we will have to consider stochastic integrals of the form $hQdW_t$ where $h \in L^p_{x,v}$, $p \geq 2$, and we should ensure the existence of the stochastic integrals as $L^p_{x,v}$-valued processes. We recall that the Lebesgue spaces $L^p_{x,v}$ with $p \geq 2$ belong to a class of the so-called 2-smooth Banach spaces, which are well suited for stochastic Ito integration (see [3], [6] for a precise construction). So, let us denote by
\( \gamma(L^2(T^N), X) \) the space of the \( \gamma \)-radonifying operators from \( L^2(T^N) \) to a 2-smooth Banach space \( X \). We recall that \( \Psi \in \gamma(L^2(T^N), X) \) if the series

\[
\sum_{k \geq 0} \gamma_k \Psi(e_k)
\]

converges in \( L^2(\Omega, X) \), for any sequence \( (\gamma_k)_{k \geq 0} \) of independent normal \( \mathcal{N}(0,1) \) random variables on a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). Then, the space \( \gamma(L^2(T^N), X) \) is endowed with the norm

\[
\|\Psi\|_{\gamma(L^2(T^N), X)} := \left( \mathbb{E} \left\| \sum_{k \geq 0} \gamma_k \Psi(e_k) \right\|_X^2 \right)^{1/2}
\]

(which does not depend on \( (\gamma_k)_{k \geq 0} \)) and is a Banach space. Now, if \( h \in L^p_{x,v}, p \geq 2 \), \( hQdW \) can be interpreted as \( \Psi dW \) where \( \Psi \) is the following \( \gamma \)-radonifying operator from \( L^2(T^N) \) to \( L^p_{x,v} \):

\[
\Psi(e_k) := hQe_k.
\]

Let us compute the \( \gamma \)-radonifying norm of \( \Psi \). We fix \( (\gamma_j)_{j \in \mathbb{N}} \) a sequence of independent \( \mathcal{N}(0,1) \)-random variables.

\[
\|\Psi\|^2_{\gamma(L^2(T^N), L^p_{x,v})} = \mathbb{E} \left\| \sum_{k} \gamma_k \Psi(e_k) \right\|_{L^p_{x,v}}^2 = \mathbb{E} \left\| \sum_{k} \gamma_k hQe_k \right\|_{L^p_{x,v}}^2
\]

\[
\leq \left( \mathbb{E} \left\| \sum_{k} \gamma_k hQe_k \right\|_{L^p_{x,v}}^p \right)^{2/p} = \left( \mathbb{E} \int_{T^N \times V} \left| \sum_{k} \gamma_k hQe_k \right|^p \right)^{2/p}.
\]

(10)

Observe that, almost everywhere in \( T^N \times V \), \( \sum_k \gamma_k hQe_k \) is a real centered Gaussian with covariance \( \sum_k |hQe_k|^2 \). As a consequence, there exists a constant \( C_p \in (0, \infty) \) such that

\[
\mathbb{E} \left\| \sum_{k} \gamma_k hQe_k \right\|_{L^p_{x,v}}^p = C_p \left( \sum_k |hQe_k|^2 \right)^{p/2}.
\]

We use this equality in the computations of the \( \gamma \)-radonifying norm to obtain, thanks to (9),

\[
\|\Psi\|^2_{\gamma(L^2(T^N), L^p_{x,v})} \leq C_p^{2/p} \left( \int_{T^N \times V} \left( \sum_k (Qe_k)^2 \right)^{p/2} |h|^p \right)^{2/p} \leq C_p^{2/p} \kappa_{0,\infty} \|h\|_{L^p_{x,v}}^2.
\]

(11)

2.3. Properties of the operator \( \sigma(\gamma)L(\cdot) \). As in the deterministic case, we expect with (1) that \( \sigma(\overline{f})L(f^\varepsilon) \) tends to zero with \( \varepsilon \), so that we should determine the equilibria of the operator \( \sigma(\gamma)L(\cdot) \). Since \( \sigma > 0 \) here, they are clearly constituted by the functions independent of \( v \in V \).

Besides, we have the following estimate: for \( f \in L^2_{x,v} \),

\[
(\sigma(\overline{f})L f, f) = -\|\sigma(\overline{f})\|^{1/2} \|Lf\|_{L^2_{x,v}}^2 \leq 0.
\]

(12)

In the space \( L^1_v \) we also have (see [1]), if \( f, g \in L^1_v \) with \( f \geq 0 \), then

\[
\int_V \text{sgn}^+(f - g) \left[ \sigma(\overline{f})L(f) - \sigma(\overline{g})L(g) \right] \, dv \leq 0,
\]

(13)
where $\text{sgn}^+(x) := 1_{x \geq 0}$. In the deterministic setting, the quantity above is involved when deriving the equation satisfied by $(f - g)^+$ where $f$ and $g$ are solutions to the equation (1) without noise and where $x^+ := \max(0, x)$ stands for the positive part of $x$. This is the main argument that permits to prove uniqueness for equation (1) without noise. In our stochastic setting, this procedure will be replaced by the application of Itô formula with the function $x \mapsto x^+$ to the process $f - g$. To make this plainly rigorous, we have to approximate the map $x \mapsto x^+$ by regular (at least $C^2$) functions. Therefore, we have to investigate what we have lost in the bound (13) above when replacing $\text{sgn}^+$ by some smooth approximation. To this end, take $\psi$ a smooth (at least $C^2$) non-decreasing function such that
\[
\begin{array}{ll}
\psi(x) = 0, & x \in (-\infty, 0], \\
\psi(x) = 1, & x \in [1, +\infty), \\
0 < \psi(x) < 1, & x \in (0, 1).
\end{array}
\]
and define
\[
\varphi_\delta(x) := \int_0^x \psi\left(\frac{y}{\delta}\right) \, dy, \quad x \in \mathbb{R}.
\] (14)
Then, we have the following lemma.

**Lemma 2.1.** Let $\delta > 0$. Suppose that $f, g \in L^1_v$ with $f \geq 0$. We have the two following estimates
\[
\int_V \varphi_\delta(f - g) \left[\sigma(\bar{f})L(f) - \sigma(\bar{g})L(g)\right] \, dv \leq C \left(1 + \|f\|_{L^1_v}\right) \delta,
\] (15)
\[
\int_V \varphi_\delta(g - f) \left[\sigma(\bar{g})L(g) - \sigma(\bar{f})L(f)\right] \, dv \leq C \left(1 + \|f\|_{L^1_v}\right) \delta.
\] (16)

**Proof.** The proof is given in Appendix A. \qed

2.4. **Main result.** We may now state our main result, the proof of which will be given throughout this paper.

**Theorem 2.2.** Let $f^\varepsilon$ denote the solution of the kinetic problem (1) in the sense of Definition 2.3 below and $\rho$ the solution of the non-linear stochastic partial differential equation
\[
\begin{aligned}
\frac{d\rho}{dt} - \text{div}_x \left(\sigma(\rho)^{-1}K \nabla_x \rho\right) &= \rho \circ Q \, dW_t, \\
\rho(0) &= \rho_{\text{in}},
\end{aligned}
\] (17)
in the sense of Definition 2.4 below, where $K$ denotes the matrix (6). Then, the solution $f^\varepsilon$ converges as $\varepsilon$ tends to 0 to the fluid limit $\rho$ and we have the estimate
\[
\sup_{t \in [0, T]} \mathbb{E}\|f_t^\varepsilon - \rho_t\|_{L^1_{x,v}} \leq C\varepsilon.
\] (18)

We refer to Section 4.6 for various remarks about the extension of the result of Theorem 2.2 to a more general framework.

In both (1) and (17) the noise is under Stratonovitch form. Actually, we interpret it, after the accurate correction, as an Itô noise. This means that the equations we are considering are in fact the kinetic equation
\[
df^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon \, dt = \frac{1}{\varepsilon^2} \sigma(\bar{f})L(f^\varepsilon) \, dt + Gf^\varepsilon \, dt + f^\varepsilon Q \, dW_t,
\]
and the fluid equation
\[
\frac{d\rho}{dt} - \text{div}_x \left(\sigma(\rho)^{-1}K \nabla_x \rho\right) = G \rho \, dt + \rho Q \, dW_t,
\]
where

\[ G := \frac{1}{2} \sum_{k \geq 0} (Qe_k)^2. \]

**Definition 2.3.** Let \( \rho_{in} \in L^2_\varepsilon \). Let \( f^\varepsilon \in L^2(\Omega; L^2(0, T; L^2_{\varepsilon,v})) \). We say that \( f^\varepsilon \) is a solution to (1) if the process \((f^\varepsilon(t))\) with values in \( L^2_{\varepsilon,v} \) is predictable and, for all \( t \in [0, T] \), and if, for all \( \varphi \in L^2_{\varepsilon,v} \) with \( \nabla_x \varphi \in L^2_{\varepsilon,v} \), \( f^\varepsilon(t) \) satisfies

\[
\langle f^\varepsilon(t), \varphi \rangle = \langle \rho_{in}, \varphi \rangle + \frac{1}{\varepsilon} \int_0^t \langle f^\varepsilon_s, a(v) \cdot \nabla_x \varphi \rangle \, ds + \frac{1}{\varepsilon} \int_0^t \langle \sigma(F^\varepsilon_s) L(f^\varepsilon_s), \varphi \rangle \, ds + \frac{1}{\varepsilon} \int_0^t \langle Gf^\varepsilon_s, \varphi \rangle \, dt + \int_0^t \langle f^\varepsilon_s, QdW_s \varphi \rangle.
\]

Note that the last term in (19) should be interpreted as

\[
\int_0^t \langle f^\varepsilon_s, QdW_s \varphi \rangle := \sum_{k \geq 0} \int_0^t \langle f^\varepsilon_s, \varphi Qe_k \rangle d\beta_k(s),
\]

an Itô integral which is well-defined by (8).

For the limit equation (17), the notion of solution that we consider is stronger:

**Definition 2.4.** Let \( \rho_{in} \in C^{2+\eta}(T^N) \) where \( \eta > 0 \). A function \( \rho \in L^2(\Omega; C^{0,2}([0, T] \times T^N)) \) is said to be a solution to (17) if the process \((\rho(t))\) with values in \( C^2(T^N) \) is predictable and, for all \( t \in [0, T] \), \( \rho(t) \) satisfies

\[
\rho(t) = \rho_{in} + \int_0^t \text{div} \left( \sigma(\rho)^{-1} K \nabla_x \rho \right) \, ds + \int_0^t G\rho \, ds + \int_0^t \rho QdW_s.
\]

**Remark 1.** The existence and uniqueness of the solution \( \rho \) with the regularity assumed in Definition 2.4 is proved in [11]. Actually, in [11] it is proved that if \( \rho_{in} \in C^{k+\eta}(T^N) \) and \( \sigma \in C^k_b \), \( k \geq 2 \), then \( \rho \in L^2(\Omega; C^{\lambda,k+\beta}([0, T] \times T^N)) \), for an exponent \( \beta > 0 \) and any \( \lambda \in (0, 1/2) \). We will use this regularity result with \( k = 3 \).

3. Formal Hilbert expansion. In this section, we derive formally the limit equation satisfied by \( f^\varepsilon \) as \( \varepsilon \) goes to 0. To do so, we classically introduce the following Hilbert expansion of the solution \( f^\varepsilon \):

\[
f^\varepsilon = f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \ldots
\]

Then, discarding the terms with positive power of \( \varepsilon \), equation (1) reads

\[
df_0 = -\frac{1}{\varepsilon} a(v) \cdot \nabla_x f_0 \, dt - a(v) \cdot \nabla_x f_1 \, dt + \frac{1}{\varepsilon} \sigma(F_0 + \varepsilon F_1 + \varepsilon^2 F_2) L(f_0 + \varepsilon f_1 + \varepsilon^2 f_2) \, dt + f_0 \circ QdW_t + O(\varepsilon).
\]

Putting the terms with the same power of \( \varepsilon \) together and omitting once again those with positive power of \( \varepsilon \), we have

\[
df_0 = \frac{1}{\varepsilon^2} \sigma(F_0) L(f_0) \, dt + \left( -\frac{1}{\varepsilon} a(v) \cdot \nabla_x f_0 + \frac{1}{\varepsilon} \sigma(F_0) L(f_1) \right. \\
+ \frac{1}{\varepsilon^2} \left[ \sigma(F_0 + \varepsilon F_1) - \sigma(F_0) \right] L(f_0) \Big) \, dt
\]

\[
+ \left( -a(v) \cdot \nabla_x f_1 + \sigma(F_0) L(f_2) + \frac{1}{\varepsilon^2} \left[ \sigma(F_0 + \varepsilon F_1 + \varepsilon^2 F_2) - \sigma(F_0 + \varepsilon F_1) \right] L(f_0) \right)
+ \frac{1}{\varepsilon} \left[ \sigma(F_0 + \varepsilon F_1) - \sigma(F_0) \right] L(f_1) \Big) \, dt + f_0 \circ QdW_t + O(\varepsilon).
\]
Next, we identify the terms in this equality with respect to the exponent of $\varepsilon$. At the order $\varepsilon^{-2}$, we find $\sigma(\mathcal{F}_0)L(f_0) = 0$, which implies $L(f_0) = 0$; thus we have $f_0 = \mathcal{F}_0 = \rho$. Then, at the order $\varepsilon^{-1}$, with the fact that $L(f_0) = 0$, we find

$$L(f_1) = \sigma(\rho)^{-1} a(v) \cdot \nabla_x \rho.$$ 

Since the integral with respect to $v \in V$ of the right-hand side vanishes thanks to (5), this equation can be solved by

$$f_1 := -\sigma(\rho)^{-1} a(v) \cdot \nabla_x \rho,$$

and we point out that $\mathcal{F}_1 = 0$. Finally, at the order $\varepsilon^0$, we get

$$d\rho = -a(v) \cdot \nabla_x f_1 \, dt + \sigma(\rho)L(f_2) \, dt + \rho \circ QdW_t.$$  \hspace{1cm} (20)

By integration with respect to $v \in V$ and with $\int_V L(f_2) dv = 0$, we discover

$$d\rho = -\text{div}_x (a(v)f_1) \, dt + \rho \circ QdW_t,$$

that is, thanks to the expression of $f_1$ given by (19),

$$d\rho - \text{div}_x (\sigma(\rho)^{-1} K \nabla_x \rho) \, dt = \rho \circ QdW_t.$$  \hspace{1cm} (21)

Furthermore, if $\rho$ satisfies equation (21), equation (20) now reads

$$\sigma(\rho)L(f_2) = \text{div}_x (\sigma(\rho)^{-1} (K - K) \nabla_x \rho),$$

where $K = a(v) \otimes a(v)$, and since the integral with respect to $v \in V$ of the right-hand side vanishes, this can indeed be solved by setting

$$f_2 := -\sigma(\rho)^{-1} \text{div}_x (\sigma(\rho)^{-1} (K - K) \nabla_x \rho).$$  \hspace{1cm} (22)

To conclude, the solution $f^\varepsilon$ of the kinetic problem (1) formally converges to an equilibrium state $\rho$ which satisfies the non-linear stochastic partial differential equation (21) given above.

4. Convergence of $f^\varepsilon$. In this section, we now give a rigorous proof of the convergence of $f^\varepsilon$. The main difficulty is that the remainder $r^\varepsilon := f^\varepsilon - \rho - \varepsilon f_1 - \varepsilon^2 f_2$ can only be appropriately estimated in $L^1_x$. As a result, in our stochastic case, we will need to apply Itô formula in $L^1_x$. This gives rise to some difficulties. So, in the sequel, we will need to push the Hilbert expansion of $f^\varepsilon$ up to order 3 to overcome these problems. To begin with, we solve the kinetic problem (1) and the limiting equation (17) and investigate the regularity and properties of the solutions.

4.1. Resolution of the kinetic problem. Let us study the kinetic problem (1). We solve it using a standard semigroup approach combined with a regularization of the random noise term. Let $p \in [1, \infty]$. We introduce the contraction semigroup $(\mathcal{U}(t))_{t \geq 0}$ generated by the linear operator $-a(v) \cdot \nabla_x$ on the space $L^p_{x,v}$.

**Proposition 1.** Let $\rho_m \in L^\infty_{x,v}$ be a non-negative function which does not depend on $v \in V$. Then there exists a unique non-negative solution $f^\varepsilon$ to the kinetic problem (1) in the sense of Definition 2.3. Furthermore, we have the following uniform bound on $f^\varepsilon$:

$$\mathbb{E} \sup_{t \in [0,T]} \|f^\varepsilon(t)\|_{L^p_{x,v}}^p \leq C,$$  \hspace{1cm} (23)

for all $p \in [1, \infty)$.

Before giving the proof of the proposition, we recall a classical result about the regularization effect of the stochastic convolution.
**Lemma 4.1.** Let \( p \in (2, \infty) \). Let \( \Psi \in L^p(\Omega; L^p(0,T; L^p_{x,v})) \). We define

\[
    z(t) := \int_0^t \mathcal{U}(t-s)\Psi(s) Q dW_s, \quad t \in [0,T].
\]

Then \( z \in L^p(\Omega; C([0,T]; L^p_{x,v})) \) and

\[
    \mathbb{E} \sup_{t \in [0,T]} ||z(t)||_{L^p_{x,v}}^p \leq C \mathbb{E} ||\Psi||_{L^p(0,T; L^p_{x,v})}^p,
\]

for some constant \( C \) which depends on \( p \) and \( \kappa_{0,\infty} \).

The proof relies on the so-called factorization method (see [7, Section 5.3]) combined with the application of the Burkholder-David-Gundy inequality for martingales with values in a 2-smooth Banach space (see [5] and [6]) and the bound (11).

**Proof of Proposition 1.** Existence and uniqueness part. In this part of the proof, for the sake of convenience, we set \( \varepsilon = 1 \).

**Step 1. Uniqueness.** We first begin with the proof of uniqueness for equation (1). So let \( f \) and \( g \) be two non-negative solutions of (1) with the same initial condition \( \rho_0 \). We set \( h := f - g \) and estimate \( h \) in \( L^1_{x,v} \) by applying the Itô formula with the \( C^2 \) function \( \varphi_\delta \) defined by (14) which approximates \( x \mapsto x^+ \). To do this we need a preliminary step of regularization: we introduce a sequence of regularizing operators \( J_r : L^2_x \rightarrow H^1_x \) which converges to the identity \( \text{Id} \) of \( L^2_x \) and is commuting with \( \nabla_x \), for example

\[
    J_r = (\text{Id} - r\Delta_x)^{-1}.
\]

With this choice \( J_r \) is also self-adjoint on \( L^2_x \). Choosing \( \varphi := J_r \varphi \) in (19), it follows from a density argument that \( J_r f \) satisfies

\[
    J_r f(t) = J_r \rho_0 - \int_0^t a(v) \cdot \nabla_x J_r f_s ds + \int_0^t J_r [\sigma(f_s) L(f_s)] ds
    + \int_0^t J_r G f_s dt + \int_0^t J_r Q dW_s.
\]

We proceed similarly with \( g \), and then apply the Itô Formula to \( J_r h = J_r f - J_r g \); this gives (note that the term relative to \( a(v) \cdot \nabla_x J_r h_s \) vanishes)

\[
    \mathbb{E} \int_{T^N \times V} \varphi_\delta(J_r h_s) = \mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta(J_r f_s - J_r g_s) J_r [\sigma(f_s) L(f_s) - \sigma(g_s) L(g_s)] ds
    + \mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta(J_r h_s) J_r [G h_s] ds
    + \mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta'(J_r h_s) G |J_r h_s|^2 ds.
\]

We take the limit \( [r \rightarrow 0] \) in this equation, to obtain

\[
    \mathbb{E} \int_{T^N \times V} \varphi_\delta(h_t) = \mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta'(f_s - g_s) [\sigma(f_s) L(f_s) - \sigma(g_s) L(g_s)] ds
    + \mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta'(h_s) G h_s ds + \mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta''(h_s) G |h_s|^2 ds.
\]
Since $x^+ \leq \varphi_\delta(x) + \delta$, we have

\[
\mathbb{E}\| (h_t)^+ \|_{L^1_{x,v}} \leq \mathbb{E} \int_{\mathbb{T}^N \times V} \varphi_\delta(h_t) + \delta.
\]

Then, for the next term, we use Lemma 2.1: we get

\[
\mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} \varphi_\delta'(f_s - g_s) \left[ \sigma(\overline{T})L(f_s) - \sigma(\overline{\sigma})L(g_s) \right] \leq C\delta \left( 1 + \mathbb{E} \int_0^T \| f_s \|_{L^1_{x,v}} ds \right)
\]

\[
\leq C\delta.
\]

For the following term, we just observe that $|\varphi_\delta'| \leq 1$ and that $\|G\|_{L^\infty_x} < \infty$ with (9) so that

\[
\mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} \varphi_\delta'(h_s)G h_s \, ds \leq C \mathbb{E} \int_0^t \| h_s \|_{L^1_{x,v}} \, ds.
\]

For the last term of the Itô formula, we point out that $\varphi_\delta''$ is zero on $[0, \delta]^c$ and that $|\varphi_\delta''| \leq C/\delta$ on $[0, \delta]$. Thus, we obtain

\[
\mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} \varphi_\delta''(h_s)G|h_s|^2 \, ds \leq C\delta.
\]

Summing up all the previous bounds now yields

\[
\mathbb{E}\| (h_t)^+ \|_{L^1_{x,v}} \leq C\delta + C \mathbb{E} \int_0^t \| h_s \|_{L^1_{x,v}} \, ds.
\]

A similar work can be done for $(h)^- = (-h)^+$. As a result we obtain the estimate

\[
\mathbb{E}\| h_t \|_{L^1_{x,v}} \leq C\delta + C \mathbb{E} \int_0^t \| h_s \|_{L^1_{x,v}} \, ds.
\]

Since this inequality holds true for all $\delta > 0$, an application of the Gronwall lemma yields $f = g$ in $L^1(\Omega; L^1(0, T; L^1_{x,v}))$.

**Step 2. Resolution of a regularized equation.** For $\delta > 0$, we will denote by $\xi_\delta$ a mollifier on $\mathbb{T}^N \times V$ as $\delta \to 0$. This step is devoted to the proof of existence of a solution $f^\delta$ to the regularized equation

\[
df + a(v) \cdot \nabla_x f \, dt = \sigma(\overline{T})L(f) \, dt + Gf \, dt + f \ast \xi_\delta QdW_t,
\]

with $\delta > 0$ being fixed. Let us fix $p > 2N$. We will apply a fixed point argument in the space $L^p(\Omega; C([0, T_0]; L^\infty_{x,v}))$ with $T_0$ sufficiently small. Before doing this, we first introduce an additional modification in the equation in order to truncate the non-linear term $f \mapsto \sigma(\overline{T})L f$, which is not globally Lipschitz. Following for example [8] or [14], we introduce $\theta \in C^\infty_0(\mathbb{R})$ whose compact support is embedded in $(-2, 2)$ and such that $\theta(x) = 1$ for $x \in [-1, 1]$ and $0 \leq \theta \leq 1$ on $\mathbb{R}$. Then, for $R > 0$, we set $\theta_R(x) = \theta(x/R)$. We are now considering the following equation:

\[
df + a(v) \cdot \nabla_x f \, dt = \theta(R(\|f\|_{L^\infty_{x,v}})\sigma(\overline{T})L(f) \, dt + Gf \, dt + f \ast \xi_\delta QdW_t,
\]

and we are looking for a mild solution $f^{R, \delta}$, that is,

\[
f(t) = \mathcal{U}(t)\rho_m + \int_0^t \mathcal{U}(t - s)\theta_R(\|f_s\|_{L^\infty_{x,v}})\sigma(\overline{T})L(f_s) \, ds + \int_0^t \mathcal{U}(t - s)Gf_s \, dt
\]

\[
+ \int_0^t \mathcal{U}(t - s)f_s \ast \xi_\delta QdW_t.
\]
Here, as usual, if \( f \in L^p(\Omega; C([0, T_0]; L^\infty_{x,v})) \), we denote by \( T f \) the right-hand side of the previous equation and we shall verify that the Banach fixed-point Theorem applies. We refer the reader to [8, Proof of Proposition 3.1] for a precise proof in a similar setting. Here, we just prove the contraction property of the stochastic integral. Thanks to Lemma 4.1 and with Young’s inequality, we easily obtain

\[
\mathbb{E} \sup_{t \in [0,T_0]} \left\| \int_0^t \mathcal{U}(t-s)(f_s - g_s) \ast \xi_s QdW_s \right\|^{p}_{L^p_{x,v}} \leq C T_0 \mathbb{E} \sup_{s \in [0,T_0]} \|f_s - g_s\|^{p}_{L^\infty_{x,v}},
\]

where the constant \( C \) depends on \( p \) and \( \kappa_{0,\infty} \). Now, since \( \nabla x \mathcal{U}(t)g = \mathcal{U}(t)\nabla v g \), we can similarly obtain

\[
\mathbb{E} \sup_{t \in [0,T_0]} \left\| \nabla_x \int_0^t \mathcal{U}(t-s)(f_s - g_s) \ast \xi_s QdW_s \right\|^{p}_{L^p_{x,v}} \leq C T_0 \mathbb{E} \sup_{s \in [0,T_0]} \|f_s - g_s\|^{p}_{L^\infty_{x,v}},
\]

where the constant \( C \) now depends on \( p, \kappa_{0,\infty}, \kappa_{1,\infty} \) and \( \|\nabla x \xi_s\|_{L^1_{x,v}} \). Furthermore, with the identity \( \nabla \mathcal{U}(t)g = -ta'(v)\mathcal{U}(t)\nabla v g + \mathcal{U}(t)\nabla v g \), a similar bound can be proved for the derivatives of the stochastic integral with respect to \( v \). To sum up, we are led to

\[
\mathbb{E} \sup_{t \in [0,T_0]} \left\| \int_0^t \mathcal{U}(t-s)(f_s - g_s) \ast \xi_s QdW_s \right\|^{p}_{W^{1,p}_{x,v}} \leq C (T_0 + T_0^2) \mathbb{E} \sup_{s \in [0,T_0]} \|f_s - g_s\|^{p}_{L^\infty_{x,v}},
\]

for some constant \( C \) which depends on \( p, \kappa_{0,\infty}, \kappa_{1,\infty}, \|\nabla x \xi_s\|_{L^1_{x,v}} \) and \( \|\nabla v \xi_s\|_{L^1_{x,v}} \).

Finally, with the Sobolev embedding \( W^{1,p}_{x,v} \subset L^\infty_{x,v} \) which holds true since \( p > 2N \), we can conclude that the contraction property of the stochastic term is satisfied in \( L^p(\Omega; C([0, T_0]; L^\infty_{x,v})) \) provided \( T_0 \) is sufficiently small. The Banach fixed-point Theorem then applies and gives us a mild solution \( f^{R,\delta}(26) \) in \( L^p(\Omega; C([0, T_0]; L^\infty_{x,v})) \).

Iterating this argument yields a solution in the space \( L^p(\Omega; C([0, T]; L^\infty_{x,v})) \). Let us introduce, for \( R > 0 \) and \( \delta > 0 \), the following stopping times

\[
\tau_{R,\delta} := \inf\{t \in [0, T], \|f^R_{\tau_{R,\delta}}\|_{L^\infty_{x,v}} > R\}.
\]

We can show, with a similar method as in [8, Lemma 4.1], that \( \tau_{R,\delta} \) is nondecreasing with \( R \) so that we can define \( \tau^*_R := \lim_{R \to \infty} \tau_{R,\delta} \). The next step is devoted to the proof of some estimates on the solution \( f^{R,\delta} \).

**Step 3. Estimates on the solution** \( f^{R,\delta} \). In this step, we emphasize the dependence through the parameters \( R \) and \( \delta \) of the constants \( C \) appearing in the estimates. For instance \( C_\delta \) depends on \( \delta \) but not on \( R \). With the mild formulation (26), using the boundedness of \( \theta_R, \sigma \) and \( G \), the contraction property of the semigroup \( \mathcal{U} \) in \( L^\infty_{x,v} \), and evaluating the stochastic integral in \( L^\infty_{x,v} \) similarly as above, we can obtain the following bound

\[
\mathbb{E} \sup_{t \in [0,T]} \|f_{t}^{R,\delta}\|_{L^p_{x,v}} \leq C_\delta, \tag{27}
\]

Note that the dependence with respect to \( \delta \) of this bound is due to the evaluation of the stochastic integral in \( L^\infty_{x,v} \), by estimating its \( W^{1,p}_{x,v} \)-norm: this gives rise to the terms \( \|\nabla x \xi_s\|_{L^1_{x,v}} \) and \( \|\nabla v \xi_s\|_{L^1_{x,v}} \) which depend on \( \delta \). Nevertheless, estimating the solution \( f^{R,\delta} \) in \( L^p_{x,v} \) with \( p > 2 \) gives a uniform bound with respect to \( R \) and \( \delta \). Precisely, with the mild formulation (26), using the boundedness of \( \theta_R, \sigma \) and \( G \), the contraction property of the semigroup \( \mathcal{U} \) in \( L^p_{x,v} \) and evaluating the stochastic...
integral in $L^p_{x,v}$, $p > 2$, thanks to Lemma 4.1, we can obtain the following bound

$$\mathbb{E} \sup_{t \in [0,T]} \| f^{R,\delta}_t \|_{L^p_{x,v}}^p \leq C. \quad (28)$$

Finally, we point out that we can also estimate $\nabla_x f^{R,\delta}_t$ in $L^p_{x,v}$, $p > 2$, by differentiating equation (26) and using the $L^\infty$ bound (27) to deal with the non-linear term. We obtain the bound

$$\mathbb{E} \sup_{t \in [0,T]} \| \nabla_x f^{R,\delta}_t \|_{L^p_{x,v}}^p \leq C. \quad (29)$$

**Step 4. Definition of $f^\delta$.** From (27) we easily deduce that for all $\delta > 0$, $\tau^*_\delta = T$ a.s. Thus, we define $f^\delta$ on $[0,T] = \cup_{R > 0} \{0, \tau_{R,\delta}\}$ by $f^\delta = f^{R,\delta}$ on $[0, \tau_{R,\delta}]$. Note that this definition makes sense since we have proved uniqueness for the equation (26) satisfied by $f^{R,\delta}$. Since $f^{R,\delta}$ is a mild solution of (25) and since for all $t \in [0,T]$ we have that $\nabla_x f^\delta$ exists a.s. in $L^p(\{0; L^p_{x,v}\})$, $p > 2$, thanks to (29) and the property $\tau^*_\delta = T$ a.s., we get that $f^\delta$ is a strong (by opposition to weak as in Definition 2.3) solution of (24), that is, $\mathbb{P}$–a.s. for all $t \in [0,T]$,

$$f^\delta(t) = \rho_{\text{in}} - \int_0^t a(v) \cdot \nabla_x f^\delta_s \, ds + \int_0^t \sigma(f^\delta_s) L(f^\delta_s) \, ds + \int_0^t G f^\delta_s \, ds + \int_0^t f^\delta_s \cdot \xi Q dW_s. \quad (30)$$

Furthermore, with (28) and the fact that $\tau^*_\delta = T$ a.s., we deduce that for $p > 2$,

$$\mathbb{E} \sup_{t \in [0,T]} \| f^\delta_t \|_{L^p_{x,v}}^p \leq C. \quad (31)$$

Finally, note that, thanks to the equation (30), we can show that $f^\delta \geq 0$. Indeed, it suffices to apply the Itô formula with the function $\varphi^\delta$ defined by (14) to the process $-f^\delta$. Similarly as in Step 1, since $\rho_{\text{in}} \geq 0$, this yields $(f^\delta)^- = 0$, hence the result.

**Step 5. Convergence $\delta \to 0$.** Thanks to (31), up to a subsequence, the sequence $(f^\delta)_{\delta > 0}$ converges weakly in $L^2(\Omega; L^2(0,T; L^2_{x,v}))$ to some $f$. This is not sufficient to pass to the limit in (30) due to the non-linear term. Thus we use the following stochastic averaging lemma, the proof of which is given in Appendix B.

**Lemma 4.2.** Let $\alpha \in (0,1]$. We assume that the non-degeneracy hypothesis (4) is satisfied. Let $f$ be bounded in $L^2(\Omega; L^2(0,T; L^2_{x,v}))$ such that

$$df + a(v) \cdot \nabla_x f \, dt = h \, dt + g \, Q \, dW_t, \quad (32)$$

with $g$ and $h$ bounded in $L^2(\Omega; L^2(0,T; L^2_{x,v}))$. Then the quantity $\rho = \overline{f}$ verifies

$$\mathbb{E} \int_0^T \| \rho_s \|^2_{H^{\alpha/2}} \, ds \leq C.$$

With (30) and (31), we apply this lemma to the process $\rho^\delta := \overline{f^\delta}$ to obtain

$$\mathbb{E} \int_0^T \| \rho^\delta_s \|^2_{H^{\alpha/2}} \, ds \leq C. \quad (33)$$

Furthermore, thanks to (30) and (31), we get that

$$\mathbb{E} \int_0^{T-h} \| f^\delta_{s+h} - f^\delta_s \|^2_{L^2_{x,v}} \, ds \leq Ch, \quad (34)$$

which also implies

$$\mathbb{E} \int_0^{T-h} \| \rho^\delta_{s+h} - \rho^\delta_s \|^2_{H^{\alpha/2}} \, ds \leq Ch. \quad (35)$$
Then, with the bounds (31) and (34) and the Aubin-Simon’s compactness lemma [20, Theorem 1] we obtain that the sequence of the laws of the processes \((f^\delta)_{\delta \geq 0}\) is tight in \(L^2(0, T; H_{x,v}^{-1})\). With the bounds (33) and (35) and [20, Theorem 4] we also get that the sequence of the laws of the processes \((\rho^\delta)_{\delta \geq 0}\) is tight in \(L^2(0, T; L^2_{x,v})\). As a consequence, with Prokhorov’s Theorem, we can assume that, up to a subsequence, the laws of the processes \((\rho^\delta)_{\delta \geq 0}\) converges weakly to the law of some process \(\rho\) in the space of probability measures on \(L^2(0, T; L^2_{x,v})\). Then, using then the Skorohod representation Theorem, there exist a new probability space \((\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})\) where lives a cylindrical Wiener process \(\hat{W}\) on the Hilbert space \(L^2(\mathbb{T}^N)\) and some random variables \(\hat{f}^\delta, \hat{f}\) with respective laws \(\mathbb{P}(f^\delta \in \cdot)\) and \(\mathbb{P}(f \in \cdot)\) such that \(\int_V f^\delta \, dv\) converges \(\hat{\mathbb{P}}\)-a.s. in \(L^2(0, T; L^2_{x,v})\) to \(\int_V \hat{f} \, dv\). Furthermore, we recall that we have the weak convergence of \(\hat{f}^\delta\) to \(\hat{f}\) in \(L^2(\hat{\Omega}; L^2(0, T; L^2_{x,v}))\). We now have all in hands to pass to the limit \(\delta \to 0\) in the weak form of (30) (obtained by testing (30) against a function \(\varphi \in L^2_{x,v}\) such that \(\nabla x\varphi \in L^2_{x,v}\)) to discover that \(\hat{\mathbb{P}}\)-a.s. for all \(t \in [0, T], (f(t), \varphi)\) satisfies (19) (with \(\varepsilon = 1\) actually).

**Step 6. Conclusion.** In this final step, we want to get rid of the change of probability space. To this purpose, we recall that we proved pathwise uniqueness for non-negative solutions to the equation (1) above in Step 1. As a consequence, we will make use of the Gyöngy-Krylov characterization of convergence in probability introduced in [15]. We recall here the precise result

**Lemma 4.3.** Let \(X\) be a Polish space equipped with the Borel \(\sigma\)-algebra. A sequence of \(X\)-valued random variables \(\{\mu_n, n \in \mathbb{N}\}\) converges in probability if and only if for every subsequence of joint laws \(\{\mu_{n_k, m_k}, k \in \mathbb{N}\}\), there exists a further subsequence which converges weakly to a probability measure \(\mu\) such that

\[
\mu((x, y) \in X \times X, x = y) = 1.
\]

Thanks to the pathwise uniqueness for non-negative solutions to (1), we can make use of this characterization of convergence in probability here (see for instance [14, Proof of Theorem 2.1] for more details about the arguments) to deduce that, up to a subsequence, the sequence \((f^\delta)_{\delta \geq 0}\) defined on the initial probability space \((\Omega, \mathcal{F}, \mathbb{P})\) converges in probability in \(L^2(0, T; L^2_{x,v})\) to a process \(f\). Without loss of generality, we can assume that the convergence is almost sure. Then, using again the method used above in Step 5, we deduce that \(\mathbb{P}\)-a.s. for all \(t \in [0, T]\),

\[
f(t) = \rho_{in} - \int_0^t a(v) \cdot \nabla x f_s \, ds + \int_0^t \sigma(\overline{f_s}) L(f_s) \, ds + \int_0^t G f_s \, dt + \int_0^t f_s Q dW_s. \tag{36}
\]

Thus \(f\) is a non-negative strong solution of the kinetic problem (1) and belongs to the expected spaces.

**Uniform bound part.** The bound (23) follows from (31).

**Remark 2.** Note that, although, by lack of a priori regularity in \(x\), we can only give a weak formulation of solution to (1) (cf. (19)), \(f^\varepsilon\) is the limit of the regularization \(f^\varepsilon\) which satisfies a mild formulation of the equation (with \(f^\varepsilon \ast \xi_{\delta Q}\) instead of \(fQ\) as a factor of the noise), cf. (30). We will use, sometimes without specifying it, this density result to justify the application of the Itô formula in some steps of the proof of the error estimate (18) below.
4.2. Definition of the two first correctors. Following the computations done in a formal way in section 3, we define:

\[
\begin{align*}
 f_1 &:= -\sigma(\rho)^{-1} a(v) \cdot \nabla_x \rho, \\
 f_2 &:= -\sigma(\rho)^{-1} \text{div}_x \left( \sigma(\rho)^{-1}(K - K) \nabla_x \rho \right).
\end{align*}
\] (37)

We state two propositions giving the properties of the processes \( f_1 \) and \( f_2 \).

Proposition 2. Let \( p \geq 1 \). The first corrector \( f_1 \), defined by (37), satisfies

\[
\sigma(\rho) L(f_1) = a(v) \cdot \nabla_x \rho
\]
with the estimate

\[
\mathbb{E} \sup_{t \in [0,T]} \| f_1(t) \|_{L^p_{x,v}}^p < \infty, \quad \mathbb{E} \sup_{t \in [0,T]} \| a(v) \cdot \nabla_x f_1(t) \|_{L^p_{x,v}}^p < \infty.
\] (39)

Furthermore, we have the following decomposition (into noise part and drift part)

\[
df_1 = f_{1,d} dt + \Psi_1 dW_t,
\] (40)

where \( \Psi_1 \) satisfies

\[
\Psi_1 = 0,
\] (41)

and where \( f_{1,d} \) and \( \Psi_1 \) satisfy

\[
\mathbb{E} \sup_{t \in [0,T]} \| f_{1,d}(t) \|_{L^p_{x,v}}^p < \infty, \quad \mathbb{E} \sup_{t \in [0,T]} \| \Psi_1(t) \|_{(L^2(T^N),L^p_{x,v})}^2 < \infty,
\] (42)

for all \( p \geq 1 \).

Proof. The equation (38) is a straightforward consequence of the definition of \( L \) and \( f_1 \) and of (5). The estimate (39) is a consequence of the regularity of \( \rho \) given in Remark 1, the bounds (H1) on \( \sigma \) and the boundedness of \( a \). By Itô’s Formula, one can easily compute the coefficient in Equation (40), in particular

\[
\Psi_1 = -a(v) \cdot \nabla_x (\sigma(\rho)^{-1} \rho Q),
\]

in the sense that

\[
\Psi_1(e_k) = -a(v) \cdot \nabla_x (\sigma(\rho)^{-1} \rho Q e_k),
\]

for all \( e_k \) element of the orthonormal basis of \( L^2(T^N) \) considered in paragraph 2.2, and

\[
f_{1,d} = a(v) \cdot \left[ \sigma'(\rho) \sigma(\rho)^{-2} \text{div}_x \left( \sigma(\rho)^{-1}(K - K) \nabla_x \rho \right) \nabla_x \rho \right] - \sigma(\rho)^{-1} a(v) \cdot \nabla_x \text{div}_x \left( \sigma(\rho)^{-1}(K - K) \nabla_x \rho \right).
\]

Consequently, we have \( \Psi_1(e_k) = 0 \) for all \( k \) since \( \pi = 0 \): this gives the identity (41). The bound (42) comes from the regularity of \( \rho \), the bounds (H1) on \( \sigma \), the regularity (H2) of \( \sigma \), the boundedness of \( a \), the bound (8) on the regularity of \( Q \) and from an estimate similar to (11) for \( \Psi_1 \).

Similarly, we can prove the following properties of the second corrector \( f_2 \):

Proposition 3. Let \( p \geq 1 \). The second corrector \( f_2 \), defined by (37), satisfies

\[
\sigma(\rho) L(f_2) = \text{div}_x \left( \sigma(\rho)^{-1}(K - K) \nabla_x \rho \right) = \text{div}_x \left( \sigma(\rho)^{-1}(K - K) \nabla_x \rho \right) + a(v) \cdot \nabla_x f_1
\] (43)

with the estimates

\[
\mathbb{E} \sup_{t \in [0,T]} \| f_2(t) \|_{L^p_{x,v}}^p < \infty, \quad \mathbb{E} \sup_{t \in [0,T]} \| a(v) \cdot \nabla_x f_2(t) \|_{L^p_{x,v}}^p < \infty.
\] (44)
Furthermore, we have the following decomposition (into noise part and drift part)
\[ df_2 = f_{2,d} \, dt + \Psi_2 \, dW_t, \]
where \( f_{2,d} \) and \( \Psi_2 \) satisfy
\[ \mathbb{E} \sup_{t \in [0,T]} \| f_{2,d}(t) \|_{L^p_{2,v}} < \infty, \quad \mathbb{E} \sup_{t \in [0,T]} \| \Psi_2(t) \|_{L^p_{2,v}} < \infty, \]
for all \( p \geq 1 \).

**Remark 3.** Later in the estimate of the remainder \( r^\varepsilon \) (see next section), we will need to decompose \( \Psi_1 \) in (40) as
\[ \Psi_1 = f_1 Q + \Psi_1^1. \]
Note then that \( \Psi_1^1 \) still satisfies (41) and (42). Note also that the process \( \Psi_1^1 \) is predictable.

4.3. **Equation satisfied by the remainder.** From now on, \( f^\varepsilon \) denotes the solution to problem (1) and \( \rho \) the solution of the limiting equation (17). We define the remainder \( r^\varepsilon \) by
\[ r^\varepsilon := f^\varepsilon - \rho - \varepsilon f_1 - \varepsilon^2 f_2 - \varepsilon^3 f_3, \]
where the correctors \( f_1, f_2 \) have been defined above. The third corrector \( f_3 \) will be defined below; its aim will be to cancel all the noise terms of order \( O(\varepsilon) \) so that the remainder has a noise term of order \( O(\varepsilon^2) \). Let us write the equation satisfied by \( r^\varepsilon \). We have
\[ dr^\varepsilon = -\frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon \, dt + \frac{1}{\varepsilon^2} \sigma(\overline{f}) L(f^\varepsilon) \, dt + f^\varepsilon \, QdW_t + G f^\varepsilon \, dt \]
\[ -d\rho - \varepsilon df_1 - \varepsilon^2 df_2 - \varepsilon^3 df_3^\varepsilon. \]
We recall that \( L(\rho) = 0 \) so that we have
\[ dr^\varepsilon = -\frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon \, dt + \frac{1}{\varepsilon} \sigma(\overline{f}) L(f^\varepsilon) \, dt + \sigma(\rho) L(f_1) \, dt + \sigma(\rho) L(f_2) \, dt + \varepsilon \sigma(\rho) L(f_3) \, dt \]
\[ + \frac{1}{\varepsilon^2} \left[ \sigma(\overline{f}) L(f^\varepsilon) - \sigma(\rho) L(f^\varepsilon - r^\varepsilon) \right] \, dt \]
\[ + f^\varepsilon \, QdW_t + G f^\varepsilon \, dt - d\rho - \varepsilon df_1 - \varepsilon^2 df_2 - \varepsilon^3 df_3^\varepsilon. \]
Using the equations satisfied by \( f_1, f_2 \) and \( \rho \), that is (38), (43) and (17), we obtain
\[ dr^\varepsilon = -\frac{1}{\varepsilon} a(v) \cdot \nabla_x f^\varepsilon \, dt + \frac{1}{\varepsilon} a(v) \cdot \nabla_x \rho \, dt + a(v) \cdot \nabla_x f_1 \, dt + \div_x \left( K \sigma(\rho)^{-1} \nabla_x \rho \right) \, dt \]
\[ + \varepsilon \sigma(\rho) L(f_3) \, dt + \frac{1}{\varepsilon^2} \left[ \sigma(\overline{f}) L(f^\varepsilon) - \sigma(\rho) L(f^\varepsilon - r^\varepsilon) \right] \, dt \]
\[ + f^\varepsilon \, QdW_t + G f^\varepsilon \, dt - \div_x \left( K \sigma(\rho)^{-1} \nabla_x \rho \right) \, dt - \rho \, QdW_t - G \rho \, dt \]
\[ - \varepsilon df_1 - \varepsilon^2 df_2 - \varepsilon^3 df_3^\varepsilon. \]
After simplification, we have,
\[ dr^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x r^\varepsilon \, dt = -\varepsilon a(v) \cdot \nabla_x f_2 \, dt - \varepsilon^2 a(v) \cdot \nabla_x f_3^\varepsilon \, dt \]
\[ + \frac{1}{\varepsilon^2} \left[ \sigma(\overline{f}) L(f^\varepsilon) - \sigma(\rho) L(f^\varepsilon - r^\varepsilon) \right] \, dt + (f^\varepsilon - \rho) \, QdW_t \]
\[ + G(f^\varepsilon - \rho) \, dt - \varepsilon df_1 - \varepsilon^2 df_2 - \varepsilon^3 df_3^\varepsilon + \varepsilon \sigma(\rho) L(f_3) \, dt. \]
Using the expression (40) of $df_1$ and Remark 3, we discover
\[
\begin{align*}
\text{dr}^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x r^\varepsilon dt &= -\varepsilon a(v) \cdot \nabla_x f_2 dt - \varepsilon^2 a(v) \cdot \nabla_x f^\varepsilon_3 dt \\
&+ \frac{1}{\varepsilon^2} [\sigma(\overline{f}) L(f^\varepsilon) - \sigma(\rho) L(f^\varepsilon - r^\varepsilon)] dt + (f^\varepsilon - \rho - \varepsilon f_1) QdW_t \\
&+ G(f^\varepsilon - \rho) dt - \varepsilon f_{1,d} dt \\
&- \varepsilon \Psi^1_r dW_t - \varepsilon^2 df_2 - \varepsilon^2 d\overline{f}^\varepsilon_3 + \varepsilon \sigma(\rho) L(f^\varepsilon_3) dt.
\end{align*}
\]

In the sequel, when estimating the remainder, we need the noise term to be of order $O(\varepsilon^2)$, see Section 4.5. As a consequence, we would like to choose $f^\varepsilon_3$ in order to compensate for the terms of order $O(\varepsilon)$ in front of the noise. Namely, we would like to impose
\[
\varepsilon^2 d\overline{f}^\varepsilon_3 - \sigma(\rho) L(f^\varepsilon_3) dt = \Psi^1_r dW_t,
\]
so that the equation satisfied by the remainder $r^\varepsilon$ is finally given by
\[
\begin{align*}
\text{dr}^\varepsilon + \frac{1}{\varepsilon} a(v) \cdot \nabla_x r^\varepsilon dt &= -\varepsilon a(v) \cdot \nabla_x f_2 dt - \varepsilon^2 a(v) \cdot \nabla_x f^\varepsilon_3 dt \\
&+ \frac{1}{\varepsilon^2} [\sigma(\overline{f}) L(f^\varepsilon) - \sigma(\rho) L(f^\varepsilon - r^\varepsilon)] dt \\
&+ (f^\varepsilon - \rho - \varepsilon f_1) QdW_t + G(f^\varepsilon - \rho) dt - \varepsilon f_{1,d} dt - \varepsilon^2 df_2.
\end{align*}
\]
Note that $f_1$ and $f_2$ do not depend on $\varepsilon$. In the following, we shall prove that $f^\varepsilon_3$ is of order $O(\varepsilon^{-1})$ with respect to $\varepsilon$. As a consequence, all the drift terms (apart from the singular one) are of order $O(\varepsilon)$. We also recall that we precisely added $f^\varepsilon_3$ in the development of $f^\varepsilon$ to get a term of order $O(\varepsilon^2)$ in front of the noise; this will be necessary further in the estimate of the remainder. We point out that $L^\varepsilon_{x,v}$ is indeed the appropriate space in which the estimate of the remainder will give a favourable sign to the singular term $\varepsilon^{-2} \left[ \sigma(\overline{f}) L(f^\varepsilon) - \sigma(\rho) L(f^\varepsilon - r^\varepsilon) \right]$, see section 2.3. Next section is devoted to the definition of the third corrector by solving equation (47).

### 4.4. Definition of the third corrector.

In this part, we study the following equation for the third corrector which was suggested in a formal way in the computations done just above:
\[
\varepsilon^2 d\overline{f}^\varepsilon_3 - \sigma(\rho) L(f^\varepsilon_3) dt = \Psi^1_r dW_t.
\]
Let us begin with some a priori estimates: if $f^\varepsilon_3$ solves (48) in the sense that
\[
f^\varepsilon_3(t) = \frac{1}{\varepsilon^2} \int_0^t \sigma(\rho(s)) L(f^\varepsilon_3(s)) ds + \frac{1}{\varepsilon^2} \int_0^t \Psi^1_r(s) dW_s
\]
for all $t \in [0,T]$ (the initial datum being 0 in particular), then, by the cancellation property (41) (and Remark 3 about $\Psi^1_r$), $\overline{f}^\varepsilon_3$ satisfies $d\overline{f}^\varepsilon_3 = 0$. Therefore $\overline{f}^\varepsilon_3(t) = 0$ for all $t$ and (49) reads
\[
f^\varepsilon_3(t) = -\frac{1}{\varepsilon^2} \int_0^t \sigma(\rho(s)) f^\varepsilon_3(s) ds + \frac{1}{\varepsilon^2} \int_0^t \Psi^1_r(s) dW_s.
\]
By Itô’s Formula, we deduce then for $p \geq 2$ that

$$
\mathbb{E}\|f_3^\varepsilon(t)\|_{L^p_{\varepsilon,v}}^p = -\frac{p}{\varepsilon^2} \mathbb{E} \int_0^t \int_{T^N \times V} \sigma(\rho(s))|f_3^\varepsilon(s)|^p ds
+ \frac{p(p-1)}{\varepsilon^2} \mathbb{E} \int_0^t \sum_k \int_{T^N \times V} |f_3^\varepsilon(s)|^{p-2} |\Psi_1^\varepsilon(e_k)|^2 ds.
$$

Using the bound from below (H1) on $\sigma$ and the bound (42) on $\Psi_1$ gives

$$
\mathbb{E}\|f_3^\varepsilon(t)\|_{L^p_{\varepsilon,v}}^p + \frac{p\sigma_*}{2\varepsilon^2} \int_0^t \mathbb{E}\|f_3^\varepsilon(s)\|_{L^p_{\varepsilon,v}}^p ds \leq C \frac{t}{\varepsilon^2},
$$

(51)

for a constant $C \geq 0$. Starting from the inequality

$$
\|f_3^\varepsilon(t)\|_{L^p_{\varepsilon,v}}^p \leq \frac{p}{\varepsilon^2} \int_0^t \sum_k \int_{T^N \times V} |f_3^\varepsilon(s)|^{p-2} f_3^\varepsilon(s)\Psi_1^\varepsilon(e_k) dW_k
+ \frac{p(p-1)}{\varepsilon^2} \int_0^t \sum_k \int_{T^N \times V} |f_3^\varepsilon(s)|^{p-2} |\Psi_1^\varepsilon(e_k)|^2 ds,
$$

and using the Burkholder-David-Gundy inequality, we also obtain an a priori estimate (depending on $\varepsilon$) on $\mathbb{E}\sup_{t \in [0,T]} \|f_3^\varepsilon(t)\|_{L^p_{\varepsilon,v}}$. With (51), these a priori estimate are enough to prove, via a Galerkin approximation scheme for example, the existence of a solution $f_3^\varepsilon$ in the sense of (49), with $f_3^\varepsilon \in C([0,T];L^p_{\varepsilon,v})$ for all $p \geq 1$, $(f_3^\varepsilon(t))$ predictable, and $f_3^\varepsilon$ satisfying the estimate (51) for real. We will show an additional estimate on $f_3^\varepsilon$ (which, again, may be rigorously justified by considering the Galerkin approximation scheme for $f_3^\varepsilon$). We apply the operator $a(v) \cdot \nabla x$ to both sides of (50) to obtain

$$
g_3^\varepsilon(t) = -\frac{1}{\varepsilon^2} \int_0^t \sigma(\rho(s)) g_3^\varepsilon(s) ds - \frac{1}{\varepsilon^2} \int_0^t \sigma'(\rho(s)) a(v) \cdot \nabla x \rho(s) f_3^\varepsilon(s) ds
+ \frac{1}{\varepsilon^2} \int_0^t \Psi_1^\varepsilon(s) dW_k,
$$

where $g_3^\varepsilon := a(v) \cdot \nabla x f_3^\varepsilon$ and $\Psi_1^\varepsilon := a(v) \cdot \nabla x \Psi_1^\varepsilon$. Note that $\Psi_1^\varepsilon$ satisfies a bound similar to the bound (42) satisfied by $\Psi_1$ and $\Psi_1^\varepsilon$. Therefore we have, as in (51), the estimate

$$
\mathbb{E}\|g_3^\varepsilon(t)\|_{L^p_{\varepsilon,v}}^p + \frac{p\sigma_*}{2\varepsilon^2} \int_0^t \mathbb{E}\|g_3^\varepsilon(s)\|_{L^p_{\varepsilon,v}}^p ds \leq C \frac{t}{\varepsilon^2} + R,
$$

where the remainder term is

$$
R = \frac{p}{\varepsilon^2} \mathbb{E} \int_0^t \int_{T^N \times V} \sigma'(\rho(s)) a(v) \cdot \nabla x \rho(s) |f_3^\varepsilon(s)|^{p-1} g_3^\varepsilon(s) ds.
$$

Since $\sigma'(\rho(s)) a(v) \cdot \nabla x \rho(s)$ is uniformly bounded, and by the Young Inequality,

$$
|f_3^\varepsilon|^{p-1} g_3^\varepsilon \leq C_{\alpha} \frac{|f_3^\varepsilon|^p}{p^\alpha} + \alpha \frac{|g_3^\varepsilon|^p}{p},
$$

for an arbitrary $\alpha > 0$, we obtain

$$
\mathbb{E}\|g_3^\varepsilon(t)\|_{L^p_{\varepsilon,v}}^p + \frac{p\sigma_*}{4\varepsilon^2} \int_0^t \mathbb{E}\|g_3^\varepsilon(s)\|_{L^p_{\varepsilon,v}}^p ds \leq \frac{C}{\varepsilon^2} + \frac{C}{\varepsilon^2} \int_0^t \mathbb{E}\|f_3^\varepsilon(s)\|_{L^p_{\varepsilon,v}}^p.
$$
Similarly, for any $\delta > \frac{484}{4\epsilon^2}$, we conclude that
\[
\mathbb{E}\|a(v) \cdot \nabla_x f^r_3(t)\|_{L^p_{x,v}}^p + \frac{ps}{4\epsilon^2} \int_0^t \mathbb{E}\|a(v) \cdot \nabla_x f^r_3(s)\|_{L^p_{x,v}}^p \, ds \leq C \frac{1}{\epsilon^2},
\]
for a given constant $C \geq 0$.

4.5. Estimate of the remainder. Finally, we estimate the remainder $r^\varepsilon$ in the space $L^1_{x,v}$; this will conclude the proof of Theorem 2.2. We point out that the correctors $f_1$, $f_2$ and $f_3$ are now properly defined in the previous sections. We recall that we set:
\[
r^\varepsilon := f^\varepsilon - \rho - \varepsilon f_1 - \varepsilon^2 f_2 - \varepsilon^3 f_3.
\]
Thanks to the calculations made in Subsection 4.3, $r^\varepsilon$ now satisfies:
\[
\frac{dr^\varepsilon}{\varepsilon} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x r^\varepsilon \, dt = -\varepsilon a(v) \cdot \nabla_x f_2 \, dt - \varepsilon^2 a(v) \cdot \nabla_x f_3 \, dt
\]
\[
+ \frac{1}{\varepsilon^2} \left[ \sigma(f^\varepsilon)L(f^\varepsilon) - \sigma(\rho)L(f^\varepsilon - r^\varepsilon) \right] \, dt
\]
\[
+ (f^\varepsilon - \rho - \varepsilon f_1) QdW_t + G(\varepsilon^r - \rho) \, dt - \varepsilon f_{1,d} \, dt - \varepsilon^2 d f_2.
\]
We will estimate $r^\varepsilon$ in $L^1_{x,v}$ by estimating $(r^\varepsilon)^+$ and $(r^\varepsilon)^-$ in $L^1_{x,v}$ using the Itô formula, where $x^+ = \max(0, x)$ and $x^- = (-x)^+$. We write the equation verified by $r^\varepsilon$ as follows:
\[
\frac{dr^\varepsilon_t}{\varepsilon} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x r^\varepsilon_t = D_t \, dt + \frac{1}{\varepsilon^2} D_t^2 \, dt + H_t \, QdW_t,
\]
where
\[
D := -\varepsilon a(v) \cdot \nabla_x f_2 - \varepsilon^2 a(v) \cdot \nabla_x f_3 + G(\varepsilon^r - \rho) - \varepsilon f_{1,d} - \varepsilon^2 f_{2,d},
\]
\[
D^\varepsilon := \sigma(f^\varepsilon)L(f^\varepsilon) - \sigma(\rho)L(f^\varepsilon - r^\varepsilon),
\]
\[
H := (f^\varepsilon - \rho - \varepsilon f_1) - \varepsilon^2 f_{2,s}.
\]
Since $f^\varepsilon - \rho = \varepsilon f_1 + \varepsilon^2 f_2 + \varepsilon^3 f_3 + r^\varepsilon$, thanks to (39), (42), (44), (46), (52) with $p = 1$ and with $\|G\|_{L^\infty} < \infty$, we have the bound:
\[
\mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} |D_s| \, ds \leq C \varepsilon + \mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} |r^\varepsilon_s| \, ds.
\]
Similarly, for any $\delta > 0$, with $f^\varepsilon - \rho - \varepsilon f_1 = \varepsilon^2 f_2 + \varepsilon^3 f_3 + r^\varepsilon_i$ and thanks to (44), (46), (51) with $p = 2$ and with $\|G\|_{L^\infty} < \infty$, we have the bound
\[
\mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} \|H_s\|^2 1_{|r^\varepsilon_s| \leq \delta} \, ds \leq C(\varepsilon^4 + \delta^2).
\]
Now, $\delta > 0$ being fixed, we apply the Itô formula with the $C^2$ approximation $\varphi_\delta$ of the function $x \mapsto x^+$ defined by (14) to the process $r^\varepsilon$ to obtain that the term relative to $\varepsilon^{-1}a(v) \cdot \nabla_x r^\varepsilon$ cancels
\[
\mathbb{E} \int_{\mathbb{T}^N \times V} \varphi_\delta(r^\varepsilon_t) = \mathbb{E} \int_{\mathbb{T}^N \times V} \varphi_\delta(r^\varepsilon_t) \, ds + \mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} \varphi_\delta'(r^\varepsilon_s) D_s \, ds \, ds
\]
\[
+ \frac{1}{\varepsilon^2} \mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} \varphi_\delta''(r^\varepsilon_s) D_s^2 \, ds + \mathbb{E} \int_0^t \int_{\mathbb{T}^N \times V} \varphi_\delta'(r^\varepsilon_s) G |H_s|^2 \, ds.
\]
Since $x^+ \leq \varphi_\delta(x) + \delta$, we have
\[
\mathbb{E}\|r^\varepsilon_t\|^2_{L^1_{x,v}} \leq \mathbb{E} \int_{\mathbb{T}^N \times V} \varphi_\delta(r^\varepsilon_t) + \delta
\]
and thanks to \( \varphi_\delta(x) \leq x^+ \), we get
\[
\mathbb{E} \int_{T^N \times V} \varphi_\delta(r^\varepsilon_{in}) \leq \mathbb{E} \|r^\varepsilon_{in}\|_{L^1_{x,v}}.
\]

With \( |\varphi_\delta'| \leq 1 \) and (53), we have
\[
\mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta'(r^\varepsilon_s) D_s \, ds \leq C \varepsilon + \mathbb{E} \int_0^t \|r^\varepsilon_s\|_{L^1_{x,v}} \, ds.
\]

Next, we study the term
\[
\int_V \varphi_\delta'(r^\varepsilon_s) D_s^2 \, dv = \int_V \varphi_\delta'(r^\varepsilon_s) \left[ \sigma(f^\varepsilon_s) L(f^\varepsilon_s) - \sigma(\rho_s) L(f^\varepsilon_s - r^\varepsilon_s) \right] \, dv.
\]

To this end, we define \( g^\varepsilon := f^\varepsilon - r^\varepsilon \); note that \( \overline{g^\varepsilon} = \rho \). The term we are interested in can be rewritten
\[
J := \int_V \varphi_\delta'(f^\varepsilon - g^\varepsilon) \left[ \sigma(\overline{f^\varepsilon}) L(f^\varepsilon) - \sigma(\overline{g^\varepsilon}) L(g^\varepsilon) \right] \, dv,
\]

so that, with the positivity of \( f^\varepsilon \), we can apply the bound (15) of Lemma 2.1 to find
\[
J \leq C(1 + \|f^\varepsilon\|_{L^1_1})\delta.
\]

We immediately deduce, using Cauchy-Schwarz’s inequality and the uniform bound (23) of \( f^\varepsilon \) in \( L^2(\Omega; L^2(0,T; L^2_{x,v})) \), that we have
\[
\frac{1}{\varepsilon^2} \mathbb{E} \int_0^t \int_{T^N \times V} \varphi_\delta'(r^\varepsilon_s) D_s^2 \, ds = \frac{1}{\varepsilon^2} \mathbb{E} \int_0^t \int_{T^N} J_s \, dx \, ds 
\leq \frac{C\delta}{\varepsilon^2} \left( 1 + \mathbb{E} \int_0^t \|f^\varepsilon_s\|_{L^2_{x,v}} \, ds \right) 
\leq \frac{C\delta}{\varepsilon^2}.
\]

Let us now study the last term of the Itô formula. We point out that \( \varphi''_\delta \) is zero on \( [0,\delta] \) and that \( |\varphi''_\delta| \leq 1/\delta \) on \( [0,\delta] \). Thus, with (54), we may write
\[
\mathbb{E} \int_0^t \int_{T^N \times V} \varphi''_\delta(r^\varepsilon_s) G[H_s]^2 \, ds \leq \frac{1}{\delta^2} \mathbb{E} \int_0^t \int_{T^N \times V} G[H_s] \, ds \leq \frac{C}{\delta} (\varepsilon^4 + \delta^2).
\]

Summing up all the previous bounds now yields
\[
\mathbb{E} \|(r^\varepsilon)^+\|_{L^1_{x,v}} \leq \mathbb{E} \|(r^\varepsilon_{in})^+\|_{L^1_{x,v}} + \delta + C\varepsilon + \mathbb{E} \int_0^t \|r^\varepsilon_s\|_{L^1_{x,v}} \, ds + \frac{C\delta}{\varepsilon^2} + \frac{C}{\delta} (\varepsilon^4 + \delta^2).
\]

Now observe that \( (r^\varepsilon)^- = (-r^\varepsilon)^+ = (g^\varepsilon - f^\varepsilon)^+ \) to obtain similarly (making use of the bound (16) instead of (15) when applying Lemma 2.1)
\[
\mathbb{E} \|(r^\varepsilon)^-\|_{L^1_{x,v}} \leq \mathbb{E} \|(r^\varepsilon_{in})^-\|_{L^1_{x,v}} + \delta + C\varepsilon + \mathbb{E} \int_0^t \|r^\varepsilon_s\|_{L^1_{x,v}} \, ds + \frac{C\delta}{\varepsilon^2} + \frac{C}{\delta} (\varepsilon^4 + \delta^2).
\]

Summing the two previous bounds and applying the Gronwall’s lemma, we get
\[
\mathbb{E} \|r^\varepsilon\|_{L^1_{x,v}} \leq C \left( \mathbb{E} \|r^\varepsilon_{in}\|_{L^1_{x,v}} + \delta + \varepsilon + \frac{\delta}{\varepsilon^2} + \frac{\varepsilon^4}{\delta} + \delta \right).
\]

Since this bound is valid for all \( \delta > 0 \), we choose \( \delta = \varepsilon^3 \) to discover
\[
\mathbb{E} \|r^\varepsilon\|_{L^1_{x,v}} \leq C \left( \mathbb{E} \|r^\varepsilon_{in}\|_{L^1_{x,v}} + \varepsilon \right).
\]
We point out that 
\[ r_\varepsilon^n = -\varepsilon f_1(0) - \varepsilon^2 f_2(0), \] 
so that 
\[ E \| r_\varepsilon \|_{L_1^x,v} \leq C\varepsilon. \]

Finally, thanks to (39), (44) and (52) with \( p = 1 \), we have
\[ \sup_{t \in [0,T]} E \| \varepsilon f_1 + \varepsilon^2 f_2 + \varepsilon^3 f_3 \|_{L_1^x} \leq C\varepsilon \]
so that we obtain the estimate
\[ \sup_{t \in [0,T]} E \| f_\varepsilon^t - \rho_t \|_{L_1^x} \leq C\varepsilon, \]
which concludes the proof of Theorem 2.2.

4.6. Final remarks.

**Remark 4** (Velocity space). Theorem 2.2 remains valid, without essential modification to the proof, if we consider a velocity space \((V,\mu)\) which is a compact Riemannian manifold without boundary with volume form \(\mu\), and a velocity functional \(a : V \to \mathbb{R}^N\) satisfying the following hypotheses.

1. \(a\) is bounded on \(V\) (cf. proof of Proposition 2 and Proposition 3),
2. there is an embedding of \(W^{1,p}(V,\mu)\) in \(L^\infty(V,\mu)\) for \(p\) large enough (cf. Step 2 in the Proof of Proposition 1),
3. \(a\) satisfies the non-stationary condition
   \[ \forall \varepsilon > 0, \forall (\xi,\sigma) \in S^{N-1} \times \mathbb{R}, \quad \mu(\{v \in V, |a(v) \cdot \xi + \sigma| < \varepsilon\}) \leq \varepsilon^\alpha, \tag{55} \]
   for some \(\alpha \in (0,1]\). Under (55), the averaging lemma Lemma 4.2 holds true with straightforward modifications.
4. the average of \(a\) over \(V\) is zero:
   \[ \int_V a d\mu = 0. \]

This is of course the equivalent to null-flux condition (5) in the case of the Torus.

The space \(V = S^{N-1}\) (unit sphere in \(\mathbb{R}^N\)) with volume form \(\mu\) and the velocity functional \(a(v) = v\), as considered in [1, 2], satisfy the items above for example.

**Remark 5** (Space Variable). We have assumed \(x \in \mathbb{T}^n\). We may consider more generally the case of \(x\) varying in an open bounded subset of \(\mathbb{R}^n\) with adequate boundary conditions, e.g. specular reflection conditions or diffusive boundary conditions, cf. [17]. Note that, at the macroscopic level, some compatibility conditions, between boundary and initial data, may be required to obtain solutions to (17) with enough regularity, [11].

**Remark 6** (Degeneracy of \(\sigma\)). One may also wonder if the error estimate of Theorem 2.2 holds true when assumption (H1) page 470 is relaxed to
\[ \sigma_* \leq \sigma(x) \leq \sigma^* + C x^{1/p} 1_{x \leq 1}, \tag{56} \]
for a given \(p > 1\) (see, e.g. [1, 2]). The answer is unknown: under such hypothesis (56), the limit equation (17) is now a degenerate (for \(\rho = 0\)) stochastic Heat equation, for which no regularity results are available.
Appendix A. Proof of Lemma 2.1.

Proof. Let us prove the first estimate; the second one is proved similarly. We are interested in the term

\[ J := \int_\mathbb{V} \varphi'_h(f - g) \left[ \sigma(\mathcal{F})L(f) - \sigma(\mathcal{G})L(g) \right] \, dv. \]

Here, we observe that

\[ 0 = \varphi'_h(\mathcal{F} - \mathcal{G}) \left[ \sigma(\mathcal{G})(\mathcal{G} - \mathcal{G}) - \sigma(\mathcal{F})(\mathcal{F} - \mathcal{F}) \right] \]

\[ = \int_\mathbb{V} \varphi'_h(\mathcal{F} - \mathcal{G}) \left[ \sigma(\mathcal{G})(\mathcal{G} - g) - \sigma(\mathcal{F})(\mathcal{F} - f) \right] \, dv. \]

As a consequence, we can write

\[ J = \int_\mathbb{V} \varphi'_h(f - g) \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{F} - \sigma(\mathcal{G})g + \sigma(\mathcal{G})g \right] \, dv \]

\[ + \int_\mathbb{V} \varphi'_h(\mathcal{F} - \mathcal{G}) \left[ \sigma(\mathcal{G})(\mathcal{G} - g) - \sigma(\mathcal{F})(\mathcal{F} - f) \right] \, dv \]

\[ = \int_\mathbb{V} \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \right] \left[ \varphi'_h(f - g) - \varphi'_h(\mathcal{F} - \mathcal{G}) \right] \, dv \]

\[ + \int_\mathbb{V} \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \right] \left[ \varphi'_h(f - g) - \varphi'_h(\mathcal{F} - g) \right] \, dv \]

\[ =: J_1 + J_2. \]

We will now bound \( J_1 \) and \( J_2 \) separately. Let us begin with the case of \( J_1 \). We decompose \( J_1 \) as:

\[ J_1 = \int_\mathbb{V} \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \right] \left[ \varphi'_h(f - g) - \varphi'_h(\mathcal{F} - \mathcal{G}) \right] 1_{f - g \leq 0} \, dv \]

\[ + \int_\mathbb{V} \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \right] \left[ \varphi'_h(f - g) - \varphi'_h(\mathcal{F} - \mathcal{G}) \right] 1_{\mathcal{F} - \mathcal{G} \leq 0} \, dv \]

\[ + \int_\mathbb{V} \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \right] \left[ \varphi'_h(f - g) - \varphi'_h(\mathcal{F} - \mathcal{G}) \right] 1_{f - g \in [0,\delta], \mathcal{F} - \mathcal{G} \in [0,\delta]} \, dv \]

\[ + \int_\mathbb{V} \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \right] \left[ \varphi'_h(f - g) - \varphi'_h(\mathcal{F} - \mathcal{G}) \right] 1_{f - g \geq \delta, \mathcal{F} - \mathcal{G} \in [0,\delta]} \, dv \]

\[ + \int_\mathbb{V} \left[ \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \right] \left[ \varphi'_h(f - g) - \varphi'_h(\mathcal{F} - \mathcal{G}) \right] 1_{f - g \geq \delta, \mathcal{F} - \mathcal{G} \geq \delta} \, dv \]

\[ =: J_1^{(1)} + J_1^{(2)} + J_1^{(3)} + J_1^{(4)} + J_1^{(5)} + J_1^{(6)}. \]

Study of \( J_1^{(1)} \): Note that when \( f - g \leq 0 \), we have \( \varphi'_h(f - g) = 0 \). If \( \mathcal{F} \leq \mathcal{G} \), we also have \( \varphi'_h(\mathcal{F} - \mathcal{G}) = 0 \), and if \( \mathcal{F} \geq \mathcal{G} \), we have \( \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \geq 0 \) thanks to the monotonicity of \( x \mapsto \sigma(x)x \) (see (H3)) and \( \varphi'_h(\mathcal{F} - \mathcal{G}) \in [0,1] \). As a result, we conclude

\[ J_1^{(1)} \leq 0. \]

Study of \( J_1^{(2)} \): Note that when \( \mathcal{F} - \mathcal{G} \leq 0 \), we have \( \varphi'_h(\mathcal{F} - \mathcal{G}) = 0 \), \( \sigma(\mathcal{F})\mathcal{F} - \sigma(\mathcal{G})\mathcal{G} \leq 0 \) thanks to the monotonicity of \( x \mapsto \sigma(x)x \) and \( \varphi'_h(f - g) \in [0,1] \) so that we obtain

\[ J_1^{(2)} \leq 0. \]
Study of $J_1^{(3)}$: First, we write

$$J_1^{(3)} = \int_V \left[ (\sigma(\overline{f}) - \sigma(\overline{g})) \overline{f} + (\sigma(\overline{g}) - f) \overline{g} \right] \left[ \varphi'_\delta(f - g) - \varphi'_\delta(\overline{f} - \overline{g}) \right] 1_{f - g \in [0,\delta], \overline{f} - \overline{g} \in [0,\delta]} \, dv.$$

Since $\varphi'_\delta(f - g) - \varphi'_\delta(\overline{f} - \overline{g}) \in [-1,1]$, we obtain with (H1) and the Lipschitz continuity of $\sigma$ (see (H2)) that

$$J_1^{(3)} \leq \int_V \left[ \|\sigma\|_{\text{Lip}} \delta + \sigma^* \delta \right] 1_{f - g \in [0,\delta], \overline{f} - \overline{g} \in [0,\delta]} \, dv \leq C(1 + |\overline{f}|)\delta.$$

Study of $J_1^{(4)}$: Note that when $\overline{f} - \overline{g} \geq \delta$ we have $\varphi'_\delta(\overline{f} - \overline{g}) = 1$ and $\sigma(\overline{f}) \overline{f} - \sigma(\overline{g}) \overline{g} \geq 0$ thanks to the monotonicity of $x \mapsto \sigma(x)x$. Since $\varphi'_\delta(f - g) \in [0,1]$, we thus get

$$J_1^{(4)} \leq 0.$$

Study of $J_1^{(5)}$: Exactly as in the case of $J_1^{(3)}$, we get

$$J_1^{(5)} \leq \int_V \left[ (\sigma(\overline{f}) - \sigma(\overline{g})) \overline{f} + (\sigma(\overline{g}) - g) \overline{g} \right] \left[ \varphi'_\delta(f - g) - \varphi'_\delta(\overline{f} - \overline{g}) \right] 1_{f - g \geq \delta, \overline{f} - \overline{g} \in [0,\delta]} \, dv \leq C(1 + |\overline{f}|)\delta.$$

Study of $J_1^{(6)}$: When $f - g \geq \delta$ and $\overline{f} - \overline{g} \geq \delta$ we have $\varphi'_\delta(f - g) = \varphi'_\delta(\overline{f} - \overline{g}) = 1$ so that

$$J_1^{(6)} = 0.$$

Now, let us study the term $J_2$. Similarly, we decompose $J_2$ as:

$$J_2 = \int_V \left[ \sigma(\overline{f})f - \sigma(\overline{g})g \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(f - g) \right] 1_{f - g \leq 0} \, dv$$

$$+ \int_V \left[ \sigma(\overline{f})f - \sigma(\overline{g})g \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(f - g) \right] 1_{f - g \geq 0} \, dv$$

$$+ \int_V \left[ \sigma(\overline{f})f - \sigma(\overline{g})g \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(\overline{f} - \overline{g}) \right] 1_{f - g \in [0,\delta], \overline{f} - \overline{g} \in [0,\delta]} \, dv$$

$$+ \int_V \left[ \sigma(\overline{f})f - \sigma(\overline{g})g \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(\overline{f} - \overline{g}) \right] 1_{f - g \in [0,\delta], \overline{f} - \overline{g} \geq \delta} \, dv$$

$$+ \int_V \left[ \sigma(\overline{f})f - \sigma(\overline{g})g \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(\overline{f} - \overline{g}) \right] 1_{f - g \geq 0, \overline{f} - \overline{g} \geq \delta} \, dv$$

$$=: J_2^{(1)} + J_2^{(2)} + J_2^{(3)} + J_2^{(4)} + J_2^{(5)}.$$

Study of $J_2^{(1)}$: When $f - g \leq 0$, we have $\varphi'_\delta(f - g) = 0$. If $\overline{f} \leq \overline{g}$, we also have $\varphi'_\delta(\overline{f} - \overline{g}) = 0$; and if $\overline{f} \geq \overline{g}$, we have $\sigma(\overline{f})f - \sigma(\overline{g})g \leq 0$ thanks to the monotonicity of $\sigma$ (see (H3)) and the positivity of $f$. Since $\varphi'_\delta(\overline{f} - \overline{g}) \in [0,1]$, we conclude

$$J_2^{(1)} \leq 0.$$

Study of $J_2^{(2)}$: When $\overline{f} - \overline{g} \leq 0$, we have $\varphi'_\delta(\overline{f} - \overline{g}) = 0$ and $\sigma(\overline{f}) - \sigma(\overline{g}) \geq 0$ thanks to the monotonicity of $\sigma$. If $f \leq \overline{g}$, we also have $\varphi'_\delta(f - g) = 0$. If $f \geq \overline{g} \geq 0$, we have $\sigma(\overline{f})f - \sigma(\overline{g})g \geq 0$. If $\overline{f} \geq \overline{g}$, we still have $\sigma(\overline{f})f - \sigma(\overline{g})g \geq 0$ since
σ ≥ 0. Note that the case 0 ≥ f ≥ g is impossible by positivity of f. Finally, since \( \varphi'_\delta(f - g) \in [0, 1] \), we conclude

\[
J_2^{(2)} ≤ 0.
\]

**Study of \( J_2^{(3)} \):** First, we write

\[
J_2^{(3)} = \int \left[ (σ(\overline{f}) - σ(\overline{g})) f + σ(\overline{g})(f - g) \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(f - g) \right] 1_{f - g ∈ [0, δ], \overline{f} - \overline{g} ∈ [0, δ]} dv.
\]

Since \( \varphi'_\delta(f - g) - \varphi'_\delta(\overline{f} - \overline{g}) ∈ [-1, 1] \), we obtain with (H1) and the Lipschitz continuity of \( σ \) that

\[
J_2^{(3)} ≤ \int \left( |f| \|σ\|_{\text{Lip}} δ + σ^* δ \right) 1_{f - g ∈ [0, δ], \overline{f} - \overline{g} ∈ [0, δ]} dv
\]

\[≤ C(1 + |\overline{f}|) δ.\]

**Study of \( J_2^{(4)} \):** We write

\[
J_2^{(4)} = \int \left[ (σ(\overline{f}) - σ(\overline{g})) f + σ(\overline{g})(f - g) \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(f - g) \right] 1_{f - g ≥ δ, \overline{f} - \overline{g} ≥ δ} dv.
\]

Note that when \( \overline{f} - \overline{g} ≥ δ \) we have \( \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(f - g) = 1 - \varphi'_\delta(f - g) ∈ [0, 1] \) and \( σ(\overline{f}) - σ(\overline{g}) ≤ 0 \) thanks to the monotonicity of \( σ \). With the positivity of \( f \), we thus get

\[
J_2^{(4)} ≤ σ^* δ.
\]

**Study of \( J_2^{(5)} \):** We have

\[
J_2^{(5)} = \int \left[ (σ(\overline{f}) - σ(\overline{g})) f + σ(\overline{g})(f - g) \right] \left[ \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(f - g) \right] 1_{f - g ∈ [0, δ], \overline{f} - \overline{g} ∈ [0, δ]} dv.
\]

Note that when \( f - g ≥ δ \) we have \( \varphi'_\delta(\overline{f} - \overline{g}) - \varphi'_\delta(f - g) = \varphi'_\delta(\overline{f} - \overline{g}) - 1 ∈ [-1, 0] \).

We thus get

\[
J_2^{(5)} ≤ \|σ\|_{\text{Lip}} \|f\|_{L^1} δ.
\]

**Study of \( J_2^{(6)} \):** When \( f - g ≥ δ \) and \( \overline{f} - \overline{g} ≥ δ \) we have \( \varphi'_\delta(f - g) = \varphi'_\delta(\overline{f} - \overline{g}) = 1 \) so that

\[
J_2^{(6)} = 0.
\]

To sum up, we get the following bound on \( J \)

\[
J ≤ C(1 + \|f\|_{L^1}) δ,
\]

which concludes the proof. \( \square \)

**Appendix B. Proof of Lemma 4.2.** We recall the Lemma to be proved.

**Lemma 4.4.** Let \( α ∈ (0, 1] \). We assume that hypothesis (4) is satisfied. Let \( f \) be bounded in \( L^2(Ω; L^2(0, T; L^2_{x,v})) \) such that

\[
df + a(v) \cdot \nabla_x f dt = h dt + g Q dW_t,
\]

with \( g \) and \( h \) bounded in \( L^2(Ω; L^2(0, T; L^2_{x,v})) \). Then the quantity \( \rho = \overline{f} \) verifies

\[
E \int_0^T \|ρ_s\|^2_{H^\alpha/2} ds ≤ C.
\]
Proof. We adapt in our stochastic context the proof of [4, Theorem 2.3]. We recall that $QdW_t = \sum_{k \geq 0} Qe_k d\beta_k(t)$ but, in order to simplify the notations, we assume in the proof that the noise is one-dimensional, namely of the form $Qe(t)$, $l \geq 0$, the generalization to an infinite dimensional noise being straightforward. We set $\theta_i = Qe_i$. Let $k \in \mathbb{Z}^N \mapsto \hat{f}(k)$ denote the Fourier transform of $f$ with respect to the space variable $x \in \mathbb{T}^N$. We take the spatial Fourier transform in Equation (57) and we add artificially on both sides of the equation a term $\lambda \hat{f}$ for some constant $\lambda > 0$ to be chosen later. We obtain, for $k \in \mathbb{Z}^N$, 

$$d\hat{f}(k) - ia(v) \cdot k \hat{f}(k) dt + \lambda \hat{f}(k) = \hat{h} dt + \hat{g_\theta} dt + \lambda \hat{f}(k).$$

Using Duhamel’s formula, we have 

$$\hat{f}(t, k) = e^{-(\lambda - ia(v) \cdot k)t} \hat{f}(0, k, v) + \int_0^t e^{-(\lambda - ia(v) \cdot k)(t - s)} \hat{h} + \lambda \hat{f}(s, k, v) ds + \int_0^t e^{-(\lambda - ia(v) \cdot k)(t - s)} \hat{g_\theta}(s, k, v) ds + \lambda \hat{f}(k).$$

Integrating in the velocity variable $v \in V$, we get 

$$\hat{\rho}(t, k) = e^{-\lambda t} \int_V e^{ia(v) \cdot k t} \hat{f}(0, k, v) dv + \int_0^t e^{-\lambda t} \int_V e^{ia(v) \cdot k (t - s)} \hat{h} + \lambda \hat{f}(s, k, v) ds dv + \int_0^t e^{-\lambda t} \int_V e^{ia(v) \cdot k (t - s)} \hat{g_\theta}(s, k, v) dv ds + \lambda \hat{f}(k).$$

where 

$$T_d(t, k) := e^{-\lambda t} \int_V e^{ia(v) \cdot k t} \hat{f}(0, k, v) dv + \int_0^t e^{-\lambda t} \int_V e^{ia(v) \cdot k (t - s)} \hat{h} + \lambda \hat{f}(s, k, v) ds dv$$

and 

$$T_s(t, k) := \int_0^t e^{-\lambda (t - s)} \int_V e^{ia(v) \cdot k (t - s)} \hat{g_\theta}(s, k, v) dv ds$$

denote respectively the deterministic and stochastic part of $\hat{\rho}(t, k)$. Let $k \in \mathbb{Z}^N$, $k \neq 0$. The deterministic term can be handled exactly as in the proof of [4, Theorem 2.3] and we obtain, up to a real multiplicative constant, 

$$E \int_0^T |T_d|^2 (t, k) dt \leq \frac{1}{\lambda^{1 - \alpha} |k|^\alpha} E \int_V |\hat{f}|^2 (0, k, v) dv + \frac{1}{\lambda^{2 - \alpha} |k|^\alpha} E \int_0^T \int_V |\hat{h} + \lambda \hat{f}|^2 (s, k, v) dv ds.$$ 

So let us now focus on the stochastic term $T_s$. First, using the Itô isometry, we have 

$$E |T_s|^2 (t, k) = E \int_0^t e^{-2\lambda (t - s)} \left| \int_V e^{ia(v) \cdot k (t - s)} \hat{g_\theta}(s, k, v) dv \right|^2 ds \leq \frac{1}{\lambda^{2 - \alpha} |k|^\alpha} E \int_0^T \int_V |\hat{h} + \lambda \hat{f}|^2 (s, k, v) dv ds.$$
so that, by the Fubini Theorem and the change of variable $\tau := t - s$, we have

$$
E \int_0^T |T_s|^2(t, k) \, dt = E \int_0^T \int_0^T e^{-2\lambda s} \left| \int_V e^{i\alpha(v) \cdot k} \tilde{g} \theta_i(\tau, k, v) \, dv \right|^2 \, ds \, d\tau
$$

$$
\leq E \int_0^T \int_0^T e^{-2\lambda s} \left| \int_V e^{i\alpha(v) \cdot k} \tilde{g} \theta_i(\tau, k, v) \, dv \right|^2 \, ds \, d\tau
$$

$$
= \frac{1}{|k|^2} E \int_0^T \int_0^T e^{-2\lambda s} \left| \int_V e^{i\alpha(v) \cdot \frac{k}{\lambda} t} \tilde{g} \theta_i(\tau, k, v) \, dv \right|^2 \, ds \, d\tau.
$$

We use the bound

$$
e^{-\frac{2\lambda s}{\alpha/k^2}} \leq \frac{1}{1 + \frac{4\lambda^2 s}{\alpha}} \quad s \geq 0,
$$

and estimate the oscillatory integral thanks to [4, Lemma 2.4] and (4); we therefore get

$$
E \int_0^T |T_s|^2(t, k) \, dt \leq \frac{C}{\lambda^{1-\alpha}|k|^\alpha} E \int_0^T \int_V |g \tilde{\theta}|^2(\tau, k, v) \, dv \, d\tau.
$$

As a result, summing up the previous bounds, we have, up to a real multiplicative constant,

$$
E \int_0^T |\tilde{\rho}|^2(t, k) \, dt \leq \frac{1}{\lambda^{1-\alpha}|k|^\alpha} E \int_0^T \int_V |g \tilde{\theta}|^2(\tau, k, v) \, dv \, d\tau
$$

$$
+ \frac{1}{\lambda^{1-\alpha}|k|^\alpha} E \int_V |\hat{f}|^2(0, k, v) \, dv
$$

$$
+ \frac{1}{\lambda^{2-\alpha}|k|^\alpha} E \int_0^T \int_V |\hat{h} + \lambda \hat{\rho}|^2(s, k, v) \, dv \, ds.
$$

We choose $\lambda \equiv 1$, multiply the last equation by $|k|^\alpha$ and sum over $k \in \mathbb{Z}^N$ to find

$$
E \int_0^T \|\rho(t)\|^2_{H^{\alpha/2}} \, dt
$$

$$
\leq CE \left[ \|g \tilde{\theta}\|_{L^2(0,T;L^2_{\alpha/2})}^2 + \|h + f\|_{L^2(0,T;L^2_{\alpha/2})}^2 + \|f(0)\|_{L^2_{\alpha/2}}^2 \right]
$$

$$
\leq CE \left[ \|Qe\|_{L^\infty_T}^2 \|g\|_{L^2(0,T;L^2_{\alpha/2})}^2 + \|h + f\|_{L^2(0,T;L^2_{\alpha/2})}^2 + \|f(0)\|_{L^2_{\alpha/2}}^2 \right].
$$

This concludes the proof when the noise is finite dimensional. For the infinite dimensional case, we recall that, with (9), we have $\kappa_{0,\infty} = \sum_{l \geq 0} \|Qe_l\|_{L^\infty_T}^2 < \infty$. \hfill \Box

**Acknowledgments.** This work is partially supported by the french government thanks to the ANR program Stosymap. It also benefit from the support of the french government “Investissements d’Avenir” program ANR-11-LABX-0020-01. Julien Vovelle was also supported by the ANR program STAB.

**REFERENCES**

[1] C. Bardos, F. Golse and B. Perthame, The Rosseland approximation for the radiative transfer equations, *Comm. Pure Appl. Math.*, **40** (1987), 691–721.

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

[3] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York, London-Sydney, 1968.

[4] P. Bouchut and L. Desvillettes, Averaging lemmas without time Fourier transform and application to discretized kinetic equations, *Proc. Roy. Soc. Edinburgh Sect. A*, **129** (1999), 19–36.
[5] Z. Brzeźniak, On stochastic convolution in Banach spaces and applications, *Stochastics Stochastics Rep.*, 61 (1997), 245–295.

[6] Z. Brzeźniak and S. Peszat, Space-time continuous solutions to SPDE’s driven by a homogeneous Wiener process, *Studia Math.*, 137 (1999), 261–299.

[7] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, 44, Cambridge University Press, Cambridge, 1992.

[8] A. de Bouard and A. Debussche, A stochastic nonlinear Schrödinger equation with multiplicative noise, *Comm. Math. Phys.*, 205 (1999), 161–181.

[9] A. de Bouard and M. Gazeau, A diffusion approximation theorem for a nonlinear PDE with application to random birefringent optical fibers, *Ann. Appl. Probab.*, 22 (2012), 2460–2504.

[10] A. Debussche and J. Vovelle, Diffusion limit for a stochastic kinetic problem, *Commun. Pure Appl. Anal.*, 11 (2012), 2305–2326.

[11] A. Debussche, S. De Moor and M. Hofmanová, A regularity result for quasilinear stochastic partial differential equations of parabolic type, *SIAM Journal on Mathematical Analysis*, 47 (2015), 1590–1614.

[12] A. Debussche, S. De Moor and J. Vovelle, Diffusion limit for the radiative transfer equation perturbed by a Markovian process, preprint, arXiv:1405.2192.

[13] J. P. Fouque, J. Garnier, G. Papanicolaou and K. Solna, *Wave Propagation and Time Reversal in Randomly Layered Media*, Stochastic Modelling and Applied Probability, 56, Springer, New York, 2007.

[14] I. Gyöngyi, Existence and uniqueness results for semilinear stochastic partial differential equations, *Stochastic Process. Appl.*, 73 (1998), 271–299.

[15] I. Gyöngyi and N. Krylov, Existence of strong solutions for Itô’s stochastic equations via approximations, *Probab. Theory Related Fields*, 105 (1996), 143–158.

[16] P.-L. Lions, B. Perthame and P. E. Souganidis, Stochastic averaging lemmas for kinetic equations, in *Séminaire Laurent Schwartz — EDP et applications (2011-2012)*, Exp. No. 26, 17pp, arXiv:1204.0317.

[17] A. Mellet and A. Vasseur, Asymptotic analysis for a Vlasov-Fokker-Planck/compressible Navier-Stokes system of equations, *Comm. Math. Phys.*, 281 (2008), 573–596.

[18] G. C. Papanicolaou, D. Stroock and S. R. S. Varadhan, Martingale approach to some limit theorems, in *Papers from the Duke Turbulence Conference* (Duke Univ., Durham, N.C., 1976), Academic Press, 1977.

[19] S. Peszat and J. Zabczyk, *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2007.

[20] J. Simon, Compact sets in the space $L^p(0; T; B)$, *Ann. Mat. Pura Appl. (4)*, 146 (1987), 65–96.

Received October 2014; revised April 2015.

E-mail address: arnaud.debussche@ens-rennes.fr
E-mail address: sylvain.demoor@ens-rennes.fr
E-mail address: vovelle@math.univ-lyon1.fr