



**HAL**  
open science

# Radiative Transfer with Long-Range Interactions: Regularity and Asymptotics

Christophe Gomez, Olivier Pinaud, Lenya Ryzhik

► **To cite this version:**

Christophe Gomez, Olivier Pinaud, Lenya Ryzhik. Radiative Transfer with Long-Range Interactions: Regularity and Asymptotics. *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, 2017, 15 (2), pp.1048 - 1072. 10.1137/15M1047076 . hal-01225235

**HAL Id: hal-01225235**

**<https://hal.science/hal-01225235>**

Submitted on 5 Nov 2015

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

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

# Radiative transfer with long-range interactions: regularity and asymptotics

Christophe Gomez, <sup>\*</sup>,<sup>1</sup>, Olivier Pinaud <sup>†</sup>,<sup>2</sup>, and Lenya Ryzhik <sup>‡</sup>,<sup>3</sup>

<sup>1</sup>Aix Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373,  
13453 Marseille, France

<sup>2</sup>Department of Mathematics, Colorado State University, Fort Collins  
CO, 80523

<sup>3</sup>Department of Mathematics, Stanford University, Stanford CA 94305

November 5, 2015

## Abstract

This work is devoted to radiative transfer equations with long-range interactions. Such equations arise in the modeling of high frequency wave propagation in random media with long-range dependence. In the regime we consider, the singular collision operator modeling the interaction between the wave and the medium is conservative, and as a consequence wavenumbers take values on the unit sphere. Our goals are to investigate the regularizing effects of grazing collisions, the diffusion limit, and the peaked forward limit. As in the case where wavenumbers take values in  $\mathbb{R}^{d+1}$ , we show that the transport operator is hypoelliptic, so that the solutions are infinitely differentiable in all variables. Using probabilistic techniques, we show as well that the diffusion limit can be carried on as in the case of a regular collision operator, and as a consequence that the diffusion coefficient is non-zero and finite. We finally consider the regime where grazing collisions are dominant.

## 1 Introduction

This work is the sequel to our first paper [17] on radiative transfer equations with long-range interactions. More precisely, we are interested in transport equations of the

---

\*christophe.gomez@univ-amu.fr

†pinaud@math.colostate.edu

‡ryzhik@math.stanford.edu

form

$$\begin{cases} \partial_t f + k \cdot \nabla_x f = \mathcal{L}f, & \text{for } (t, x, k) \in \mathbb{R}_+ \times \mathbb{R}^{d+1} \times \mathbb{S}^d \\ f(t=0, \cdot, \cdot) = f^0 \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d) \\ \mathcal{L}f(k) = \int_{\mathbb{S}^d} F(k \cdot p) (f(p) - f(k)) d\sigma(p). \end{cases} \quad (1)$$

Above,  $d \geq 1$ ,  $\mathbb{S}^d$  is the unit sphere in  $\mathbb{R}^{d+1}$  and  $d\sigma(p)$  is the surface measure on  $\mathbb{S}^d$ . Here, the collision kernel  $F$  is non-negative and non-integrable at zero so that

$$\int_{\mathbb{S}^d} F(k \cdot p) d\sigma(p) = \infty.$$

Our motivation for studying such equations stems from the analysis of wave propagation in random media with long-range dependence. For high frequency waves in the weak coupling regime, the wave energy is asymptotically described by a transport equation of the form (1) with a singular kernel. The fact that the collision process is confined to the sphere is due to the conservative nature of the interaction between the wave and the random medium. Other types of collision kernel arise depending on the scalings. The one we consider has the form

$$\mathcal{L}f(k) = \int_{\mathbb{R}^{d+1}} \delta\left(\frac{|k|^2}{2} - \frac{|p|^2}{2}\right) \hat{R}(k-p)(f(p) - f(k)) dp,$$

where  $\delta$  is the Dirac measure, that decouples the transport equations for different values of  $|k|$ , and we may set  $|k| = 1$ . The term  $\hat{R}$  is the power spectrum of a random potential  $V$ , mean-zero and statistically homogeneous in space,

$$R(x) = \mathbb{E}[V(x+y)V(y)] = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^{d+1}} e^{ik \cdot x} \hat{R}(k) dk,$$

that models random fluctuations in the Schrödinger equation

$$i\partial_t \psi^\varepsilon + \frac{1}{2} \Delta_x \psi^\varepsilon - \sqrt{\varepsilon} V(x) \psi^\varepsilon = 0.$$

Here,  $\varepsilon \ll 1$  is the variance of the fluctuations. Equation (1) then describes energy transport in a certain macroscopic limit via asymptotics of the Wigner transform [15, 23, 30]. See [7, 10, 24, 28, 16] for more details on the link between (1) and wave propagation in random media.

In the present work, we investigate the regularity of the solutions of (1) and two asymptotic regimes. We show that the transport operator is hypoelliptic, namely that for any initial condition in  $L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , the solution is  $\mathcal{C}^\infty$  in all variables for positive times. This is a consequence of the singular nature of the collision operator when the correlation function  $R(x)$  decays only algebraically:

$$\text{when } R(x) \sim \frac{1}{|x|^{2-2\alpha}}, \quad |x| \gg 1, \quad \text{then } \hat{R}(p) \sim \frac{1}{|p|^{2\alpha+d-1}},$$

and this singularity is non-integrable when  $\alpha \in (1/2, 1)$  (recall that the integration is carried over the  $d$ -dimensional sphere). Therefore, we have (as  $|k| = |p| = 1$ )

$$\hat{R}(k-p) = \hat{R}(|k-p|) = \hat{R}(2\sqrt{1-\cos\theta}) \sim \theta^{-d-2\beta}, \quad \text{as } \theta \rightarrow 0,$$

where  $\beta = 2\alpha - 1 \in (0, 1)$ . The regularizing effect of grazing collisions is now well-established in the case momenta take values in  $\mathbb{R}^{d+1}$  and not just on the sphere, see e.g. [4, 18, 19, 3, 1, 2, 9, 22] for references on the Boltzmann equation in the so-called non cut-off case. The heuristics of the regularization effect is similar to the case of collisions over  $\mathbb{R}^{d+1}$ , and goes as follows: the high frequency behavior of the collision operator  $\mathcal{L}$  is essentially that of a fractional Laplace-Beltrami operator on the sphere, and as a consequence standard energy estimates yield some Sobolev regularity in the  $k$  variable. This regularity is then propagated to the other variables by the free transport operator. The proof of regularity is based on the following hypoelliptic estimates that we derived in [17]:

**Theorem 1.1** *Assume  $g, h \in L^2(\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d)$ , and let  $f \in L^2(\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d)$  satisfy the transport equation*

$$\partial_t f + k \cdot \nabla_x f = (-\Delta_d)^\beta h + g \quad (2)$$

*in the distribution sense, where  $\Delta_d$  is the Laplace-Beltrami operator on  $\mathbb{S}^d$  and  $\beta \geq 0$ . For some  $\theta > 0$ , suppose, in addition, that*

$$(-\Delta_d)^{\frac{\theta}{2}} f \in L^2(\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d).$$

*Then, for*

$$\gamma = \frac{\theta}{2(1 + 2\beta) + \theta},$$

*we have  $\partial_{t,x}^\gamma f \in L^2(\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d)$  with the estimate*

$$\|\partial_{t,x}^\gamma f\|_{L^2} \leq C \left( \|(-\Delta_d)^{\frac{\theta}{2}} f\|_{L^2} + \|f\|_{L^2} + \|g\|_{L^2} + \|h\|_{L^2} \right).$$

The fractional derivatives above are defined in the Fourier space, see the notation section further. Note that we stated here a slightly more general version of the theorem of [17] that better suits our needs for this work: we added an additional source term  $g$  in (2) which essentially does not modify the original proof. The proof of Theorem 1.1 follows the techniques of Bouchut [9], that are adapted to the spherical geometry. Replacing formally  $(-\Delta_d)^\beta h$  in the theorem by  $\mathcal{L}f$ , it is expected, using bootstrapping arguments, that  $f$  has derivatives of any order in all variables.

We will also address two asymptotic regimes. The first one is the diffusion regime where the long time behavior of solutions to (1) is investigated. The main question is whether the singular nature of the collision operator leads to a different equation than the standard diffusion equation obtained for smooth collision kernels. The answer is no with the singularities that we consider, we obtain a perfectly defined diffusion matrix. The second regime is the peaked forward regime where grazing collisions are the most significant. In the case where collision are defined over  $\mathbb{R}^{d+1}$ , this leads to the well-known Fokker-Planck (or fractional Fokker-Planck) equation. Here, the situation is slightly different, the limiting collision operator is not a fractional Laplace-Beltrami operator, but an operator that shares the same high frequency behavior as the Laplace-Beltrami, but not the same low frequency behavior. We will base our asymptotic analysis on the probabilistic representation of the transport equation (1). This is mostly motivated by (future) numerical considerations: while standard discretizations of (1), using e.g. finite

elements of finite volumes, might not be straightforward due to the singularity of the kernel, probabilistic methods offer a simple alternative. The operator  $\mathcal{L}$  can be seen as the generator of a jump Markov process, which can be easily simulated. The solution to (1) is then obtained after averaging over several realizations of the process.

The paper is structured as follows: in Section 2, we introduce the notation and state the hypotheses on the kernel  $F$ . We present our main results in Section 3. Theorem 3.1 concerns the regularity theory, Theorem 3.2 the diffusion limit, and Theorem 3.3 the peaked forward limit. The probabilistic representation of the solution to (1) is introduced before Theorem 3.2. Section 4 is devoted to the proof of Theorem 3.1, and includes important results on the operator  $\mathcal{L}$  that will be used in the proofs of Theorems 3.2 and 3.3, given in Sections 5 and 6, respectively.

To conclude this introduction, we would like to mention, that after this second work on radiative transfer equations with long-range interactions was completed, we became aware of the recent paper [5] that addresses similar questions. Some differences are the following: it seems our hypoelliptic estimates of Theorem 1.1 are sharper; our techniques of proof are different, we use in particular probabilistic techniques for the peaked forward regime instead of PDE techniques; the diffusion limit is not addressed in [5].

**Acknowledgment.** This work was supported in part by NSF grant DMS-1311903, an AFOSR NSSEFF Fellowship, and NSF CAREER grant DMS-1452349.

## 2 Preliminaries

We introduce here some notations and the main properties of the collision kernel  $F$ .

**Notation.** We will denote by  $\langle \cdot, \cdot \rangle$  the  $L^2(\mathbb{S}^d)$  inner product. The Laplace-Beltrami operator  $\Delta_d$  is defined by

$$\begin{aligned} \Delta_d \varphi(z) &= \Delta \varphi \left( \frac{y}{\|y\|} \right) \Big|_{y=z}, \quad z \in \mathbb{S}^d \\ &= \left( \Delta - \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} z_i z_j \partial_{z_i} \partial_{z_j} - d \sum_{i=1}^{d+1} z_i \partial_{z_i} \right) \varphi(z). \end{aligned}$$

Here,  $\Delta$  is the standard Laplacian on  $\mathbb{R}^{d+1}$ . The eigenfunctions of  $-\Delta_d$  are the spherical harmonics  $Y_{n,m}$ , for  $n \in \mathbb{N}$  and  $m = 1, \dots, M(d, n)$  where

$$M(d, n) = (2n + d - 1) \frac{\Gamma(n + d - 1)}{\Gamma(d) \Gamma(n + 1)},$$

and are associated with the eigenvalues  $\lambda_n = n(n + d - 1)$ . Above,  $\Gamma$  is the gamma function. We will denote by  $e_0 = 1/(\sigma(\mathbb{S}^d))^{1/2}$  the first spherical harmonics  $Y_{0,1}$ . The Fourier representation of the fractional Laplace-Beltrami operator is then, for  $\beta \in (0, \infty)$ ,

$$(-\Delta_d)^\beta \varphi(k) = \sum_{n=0}^{\infty} \sum_{m=1}^{M(d,n)} \lambda_n^\beta \langle \varphi, Y_{n,m} \rangle Y_{n,m}(k),$$

where convergence is understood in  $L^2(\mathbb{S}^d)$ . The Sobolev space  $H^\theta(\mathbb{S}^d)$ , for  $\theta > 0$ , is defined by

$$H^\theta(\mathbb{S}^d) = \left\{ \varphi \in L^2(\mathbb{S}^d), \quad (-\Delta_d)^{\frac{\theta}{2}} \varphi \in L^2(\mathbb{S}^d) \right\},$$

and is equipped with the norm

$$\|\varphi\|_{H^\theta(\mathbb{S}^d)}^2 = \|\varphi\|_{L^2(\mathbb{S}^d)}^2 + \|(-\Delta_d)^{\frac{\theta}{2}} \varphi\|_{L^2(\mathbb{S}^d)}^2.$$

We will denote by  $\mathcal{C}_c^p$  the space of  $\mathcal{C}^p$  functions with compact support, and by  $\mathcal{C}_b^p$  the space of bounded  $\mathcal{C}^p$  functions. We will use the following convention for the Fourier transform in  $\mathbb{R}^{d+1}$ :

$$\hat{f}(\xi) = \mathcal{F}f(\xi) = \int_{\mathbb{R}^{d+1}} e^{-ix \cdot \xi} f(x) dx, \quad \mathcal{F}^{-1}f(x) = \frac{1}{(2\pi)^{d+1}} \int_{\mathbb{R}^{d+1}} e^{ix \cdot \xi} \hat{f}(\xi) d\xi,$$

and introduce the fractional derivative as

$$\partial_{x_j}^\gamma f(x) = \mathcal{F}^{-1}[(i\xi_j)^\gamma \hat{f}(\xi)](x),$$

with a similar definition for fractional derivatives involving the time variable. We will denote by  $\partial_{t,x}^\gamma f$  any of the fractional derivatives with respect to  $t$  or  $x_j$ .

**The collision kernel.** We suppose  $F$  is non-negative and satisfies the following hypotheses:

$$\forall \delta \in (-1, 1), \quad F(s)(1 - s^2)^{\frac{d-2}{2}} \in L^1((-1, \delta)),$$

and there exists  $a_1 \in (0, \infty)$  such that

$$\lim_{s \rightarrow 1} |1 - s|^{\beta + \frac{d}{2}} F(s) = a_1, \quad \beta \in (0, 1). \quad (3)$$

We will use the following decomposition of the kernel into smooth and singular parts: let first

$$a_2(s) = \frac{1}{a_1} (|1 - s|^{\beta + \frac{d}{2}} F(s) - a_1).$$

According to (3),  $a_2(s) \rightarrow 0$  as  $s \rightarrow 1$ , and therefore there exists  $\delta \in (0, 1)$  such that  $|a_2(s)| < 1$  for  $s \in (\delta, 1)$ . Let then  $\chi \in \mathcal{C}^\infty([-1, 1])$  such that  $\chi(s) = 1$  for  $s \in [-1, \delta]$ , and  $\chi(s) = 0$  for  $s \in [\delta', 1]$  with  $\delta' \in (\delta, 1)$ . The kernel is finally written as

$$F(s) = \chi(s)F(s) + (1 - \chi(s)) \left( \frac{a_1(1 + a_2(s))}{|1 - s|^{\beta + \frac{d}{2}}} \right) := F_1(s) + F_2(s). \quad (4)$$

We have by construction  $F_1 \in L^1((-1, 1))$ .

Note that the so-called mean free path, defined as the inverse of the integral of  $F$ , is equal to zero since

$$\int_{\mathbb{S}^d} F(k \cdot p) d\sigma(p) \sim \int_{\mathbb{S}^d} \frac{d\sigma(p)}{|1 - k \cdot p|^{\beta + \frac{d}{2}}} \sim \int_{-1}^1 \frac{(1 - t^2)^{\frac{d-2}{2}}}{|1 - t|^{\beta + \frac{d}{2}}} dt = \infty.$$

For  $f \in \mathcal{C}^\infty(\mathbb{S}^d)$  (such a regularity is in fact not needed), the collision operator is defined more rigorously by

$$\mathcal{L}f(k) = \text{p.v.} \int_{\mathbb{S}^d} F(k \cdot p) (f(p) - f(k)) d\sigma(p),$$

where p.v. stands for the principal value. It is shown in Lemma 4.1 that  $\mathcal{L}f \in L^\infty(\mathbb{S}^d)$  for such  $f$ .

### 3 Main results

We present two types of result: our first result concerns the regularity of the solutions to (1), and our other results address the asymptotic behavior of the solutions in two different regimes, the diffusion regime, and the peaked forward regime. We start with the regularity results.

**Hypoelliptic estimates and regularity.** For any  $f_0 \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , we say that  $f \in L^\infty((0, \infty), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))$  is a weak solution to (1), if, for all  $\varphi \in \mathcal{C}^\infty([0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  with compact support in  $x$ , with in addition  $\varphi = 0$  for  $t \geq T$ ,  $T$  arbitrary,

$$\begin{aligned} & \int_0^\infty \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f(t, x, k) (\partial_t + k \cdot \nabla_x + \mathcal{L}) \varphi(t, x, k) dt dx d\sigma(k) \\ &= - \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f^0(x, k) \varphi(0, x, k) dx d\sigma(k). \end{aligned}$$

Note that the definition makes sense since  $\mathcal{L}\varphi \in L^\infty(\mathbb{S}^d)$  when  $\varphi \in \mathcal{C}^\infty(\mathbb{S}^d)$  according to Lemma 4.1. Our first theorem is the following:

**Theorem 3.1** (*Regularity*) *The operator  $\partial_t + k \cdot \nabla_x - \mathcal{L}$  is hypoelliptic. Namely, for any  $f^0 \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , (1) admits a unique weak solution  $f$  that satisfies*

$$f \in \mathcal{C}^\infty((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d).$$

The proof of Theorem 3.1 is given in Section 4. It is based on a regularization procedure in order to obtain, along with uniform estimates, existence and uniqueness of weak solutions. We then use Theorem 1.1 in order to gain better regularity in the  $(t, x)$  variables first, and show in a second step the improved regularity in the  $k$  variable.

**Probabilistic representation and asymptotics.** We give below a probabilistic interpretation of the solutions to (1), and perform the asymptotic analysis in the probabilistic framework. The starting point is the fact, proved in Lemma 4.1 further, that the operator  $\mathcal{L}$  defined on  $\mathcal{C}^\infty$  can be extended to a unique non-positive self-adjoint operator  $\bar{\mathcal{L}}$  with domain  $\mathcal{D}(\bar{\mathcal{L}}) = \{\varphi \in H^\beta(\mathbb{S}^d), \bar{\mathcal{L}}\varphi \in L^2(\mathbb{S}^d)\}$ . Moreover,  $-\bar{\mathcal{L}}$  is associated to a quadratic form  $\mathcal{Q} : H^\beta(\mathbb{S}^d) \times H^\beta(\mathbb{S}^d) \rightarrow \mathbb{R}_+$  given by

$$\mathcal{Q}(f, g) = \frac{1}{2} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} F(k \cdot p) (f(k) - f(p))(g(k) - g(p)) d\sigma(k) d\sigma(p).$$

Then, according to [13, Theorem 7.2.1 pp. 302], there exists a Markov process

$$\mathbf{M} = (\Omega, \mathcal{F}, (m(t))_{t \geq 0}, (\mathbb{P}_k)_{k \in \mathbb{S}^d})$$

on  $\mathbb{S}^d$ , càdlàg (right continuous with left limits), with generator  $\bar{\mathcal{L}}$ , and where  $(\Omega, \mathcal{F})$  is a measurable space. The subscript  $k \in \mathbb{S}^d$  in the measure  $\mathbb{P}_k$  is the initial condition of the Markov process. For  $x \in \mathbb{R}^{d+1}$ , considering then

$$X_x(t) = \left( x - \int_0^t m(u) du, m(t) \right),$$

we have a Markov process on  $\mathbb{R}^{d+1} \times \mathbb{S}^d$  with generator

$$\tilde{\mathcal{L}}\varphi = -k \cdot \nabla_x \varphi + \bar{\mathcal{L}}\varphi, \quad \text{whose domain includes } \mathcal{C}^1(\mathbb{R}^{d+1}, \mathcal{D}(\bar{\mathcal{L}})).$$

Introducing  $T_t$  the semigroup  $T_t f^0(x, k) = \mathbb{E}_k[f^0(X_x(t))]$  associated to  $\tilde{\mathcal{L}}$ , where  $\mathbb{E}_k$  is the expectation with respect to the measure  $\mathbb{P}_k$ , it follows, for instance from [11, Proposition 1.5 pp. 9], that for all  $f^0 \in \mathcal{C}^1(\mathbb{R}^{d+1}, \mathcal{D}(\bar{\mathcal{L}})) \cap L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , the function  $u(t, x, k) := T_t f^0(x, k) \in \mathcal{C}^1(\mathbb{R}^{d+1}, \mathcal{D}(\bar{\mathcal{L}}))$  satisfies,

$$\partial_t u = \tilde{\mathcal{L}}u, \quad \text{for all } (t, x, k) \in (0, +\infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d.$$

Above and in the sequel, we use the shorthand  $f^0(X_x(t))$  for  $f^0((X_x(t))_1, (X_x(t))_2)$ . According to Theorem 3.1, the latter equation admits a unique smooth solution, and therefore  $u(t, x, k) = f(t, x, k)$ . The semigroup  $T_t$  can be extended to  $L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$  thanks to the following estimate, proved in Proposition 4.2,

$$\|f\|_{L^\infty((0, \infty), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))} \leq \|f_0\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}.$$

In this probabilistic framework, we therefore have

$$f(t, x, k) = T_t f^0(x, k) = \mathbb{E}_k[f^0(X_x(t))],$$

and we investigate now two asymptotic regimes. The first one is the diffusion regime, where strong collision effects are investigated in the long time limit. There are two equivalent ways to study the limit, either by rescaling the collision operator as  $\mathcal{L} \rightarrow \mathcal{L}/\varepsilon$  and the time variable as  $t \rightarrow t/\varepsilon$ , or by the change of variables (this is the route we follow here)

$$t \rightarrow \frac{t}{\varepsilon^2}, \quad x \rightarrow \frac{x}{\varepsilon}, \quad f(t, x, k) \rightarrow f^\varepsilon(t, x, k) = f\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}, k\right),$$

with  $f^\varepsilon(0, x, k) = f^0(x, k)$ . The transport equation (1) then becomes

$$\varepsilon^2 \partial_t f^\varepsilon + \varepsilon k \cdot \nabla_x f^\varepsilon = \mathcal{L} f^\varepsilon. \quad (5)$$

In this regime, the probabilistic representation is obtained by considering the rescaled Markov process

$$X_x^\varepsilon(t) = \left( x - \varepsilon \int_0^{t/\varepsilon^2} m(u) du, m(t/\varepsilon^2) \right),$$



with generator

$$\tilde{\mathcal{L}}_\varepsilon \varphi = -\frac{1}{\varepsilon} k \cdot \nabla_x \varphi + \frac{1}{\varepsilon^2} \bar{\mathcal{L}} \varphi.$$

The solution to (5) then reads

$$f^\varepsilon(t, x, k) = T_t^\varepsilon f^0(x, k) = \mathbb{E}_k[f^0(X_x^\varepsilon(t))].$$

Our second result is the following:

**Theorem 3.2** (*Diffusion limit*) *Suppose there exists  $\delta > 0$  such that  $F \geq \delta$  a.e. on  $(-1, 1)$ , and let  $Y_x^\varepsilon = x - \varepsilon \int_0^{t/\varepsilon^2} m(u) du$ . Then, for all  $x \in \mathbb{R}^{d+1}$ , the process  $Y_x^\varepsilon$  converges in law in  $\mathcal{C}^0([0, \infty), \mathbb{R}^{d+1})$  to a diffusion process  $Y_x$ , starting at  $x$ , with generator*

$$\tilde{\mathcal{L}}_0 = \nabla_x \cdot D \nabla_x,$$

where the positive-definite diffusion matrix  $D$  is given by

$$D_{jl} = \frac{1}{\mathfrak{C}} \int_{\mathbb{S}^d} d\tilde{\sigma}(k) k_j k_l \quad \text{with} \quad \mathfrak{C} = \sigma(\mathbb{S}^{d-1}) \int_{-1}^1 F(s) (1-s^2)^{(d-2)/2} (1-s) ds.$$

Above,  $\tilde{\sigma}$  is the uniform measure on  $\mathbb{S}^d$  (i.e.  $\tilde{\sigma} = \sigma/\sigma(\mathbb{S}^d)$ ). Moreover, for any  $f^0 \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$  and for all  $t > 0$ ,  $f_\varepsilon(t)$  converges weakly in  $L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$  to the unique solution to

$$\partial_t f = \nabla_x \cdot (D \nabla_x f) \quad \text{with} \quad f(0, x) = \int_{\mathbb{S}^d} d\tilde{\sigma}(p) f^0(x, p), \quad (6)$$

and the function  $f$  reads

$$f(t, x) = \mathbb{E}^{Y_x} [f(0, y_t)]$$

where  $\mathbb{E}^{Y_x}$  denotes expectation with respect to the law of  $Y_x$ , and  $y$  is the canonical map defined by  $y_t(\omega) = \omega(t)$  for all  $\omega \in \mathcal{C}^0([0, \infty), \mathbb{R}^{d+1})$ .

Let us give a few comments on this theorem. The hypothesis that  $F$  is strictly positive is needed for the spectral gap estimate of Lemma 4.1. Moreover, the constant  $\mathfrak{C}$  is non-zero and finite, and so does  $D$  since, according to (3),

$$F(s)(1-s^2)^{(d-2)/2}(1-s) \sim (1-s)^{-\beta} \in L^1((-1, 1)) \quad \text{for} \quad \beta \in (0, 1).$$

Finally, probabilistic techniques yield slightly stronger convergence results than standard techniques that would provide weak- $\star$  convergence in  $L^2((0, T) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . Here, convergence is pointwise in  $t$  and weak in  $L^2$  for  $L^2$  initial conditions. Naturally, convergence can be made stronger by introducing the first order corrector, we did not pursue this route here.

The main ingredient in the proof (given in Section 5) is a spectral gap estimate that shows that the Markov process  $Y_x^\varepsilon$  is ergodic, which allows us to use standard diffusion-approximation theorems to obtain convergence.

Our last result concerns the peaked forward limit. In this regime, the main contribution to the scattering process is due to grazing collisions. This is translated to the

kernel  $F$  by supposing that  $F(s)$  is small for  $s \neq 1$ , and that  $F(1)$  is large. According to (3), this suggests the scaling

$$F^\varepsilon(s) = \varepsilon^{\beta+d/2} K(\varepsilon(1-s)), \quad \text{with} \quad K(t) \underset{t \rightarrow 0}{\sim} \frac{a_1}{|t|^{\beta+d/2}}, \quad (7)$$

and  $(1-s^2)^{\frac{d-2}{2}} K(1-s) \in L^1((-1, \delta))$ , for all  $\delta \in (-1, 1)$ . The rescaled generator is accordingly  $\tilde{\mathcal{L}}_\varepsilon$ , and the corresponding Markov process is now

$$\mathbf{M}^\varepsilon = (\Omega, \mathcal{F}, (m^\varepsilon(t))_{t \geq 0}, (\mathbb{P}_k)_{k \in \mathbb{S}^d}).$$

Consider then

$$X_x^\varepsilon(t) = \left( x - \int_0^t m^\varepsilon(u) du, m^\varepsilon(t) \right),$$

with generator

$$\tilde{\mathcal{L}}_\varepsilon \varphi = -k \cdot \nabla_x \varphi + \bar{\mathcal{L}}_\varepsilon \varphi,$$

and associated semigroup  $T_t^\varepsilon$ . Then, the function  $f^\varepsilon(t, x, k) = T_t^\varepsilon f^0(x, k) = \mathbb{E}_k[f^0(X_x^\varepsilon(t))]$  satisfies

$$\partial_t f^\varepsilon + k \cdot \nabla_x f^\varepsilon = \bar{\mathcal{L}}_\varepsilon f^\varepsilon. \quad (8)$$

Above,  $\mathbb{E}_k$  denotes expectation with respect to the measure  $\mathbb{P}_k$ . Let us denote by  $\mathbb{X}_{x,k}^\varepsilon$  the law of  $X_x^\varepsilon$  starting at  $(x, k)$ . Denoting by  $\mathcal{D}([0, \infty))$  the space of càdlàg functions equipped with the Skorohod topology [8], our third result is the following:

**Theorem 3.3** (*Peaked forward limit*) *For all  $(x, k) \in \mathbb{R}^{d+1} \times \mathbb{S}^d$ , the measure  $\mathbb{X}_{x,k}^\varepsilon$  converges weakly in  $\mathcal{D}([0, \infty), \mathbb{R}^{d+1} \times \mathbb{S}^d)$  to a measure  $\mathbb{X}_{x,k}$ , which is the law of a diffusion process starting at  $(x, k)$  with generator*

$$\tilde{\mathcal{L}}_0 = -k \cdot \nabla_x + \mathcal{L}_\beta,$$

where

$$\mathcal{L}_\beta \varphi(k) = p.v. a_1 \int_{\mathbb{S}^d} \frac{\varphi(p) - \varphi(k)}{|p - k|^{2\beta+d}} d\sigma(p).$$

Moreover, for any  $f^0 \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$  and all  $t > 0$ ,  $f^\varepsilon(t)$  converges weakly in  $L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$  to the unique solution to

$$\partial_t f + k \cdot \nabla_x f = \mathcal{L}_\beta f \quad \text{with} \quad f(0, x, k) = f^0(x, k), \quad (9)$$

and the function  $f$  reads

$$f(t, x, k) = \mathbb{E}^{\mathbb{X}_{x,k}}[f^0(y_t)],$$

where  $\mathbb{E}^{\mathbb{X}_{x,k}}$  denotes the expectation with respect to  $\mathbb{X}_{x,k}$ , and  $y$  is the canonical map defined by  $y_t(\omega) = \omega(t)$  for all  $\omega \in \mathcal{D}([0, \infty), \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . When  $f^0 \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d) \cap \mathcal{C}_b^0(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , then the convergence holds pointwise in  $(t, x, k)$ .

Note that the operator  $\mathcal{L}_\beta$  is not the fractional Laplace-Beltrami operator. It would be in the Euclidean case if  $\mathbb{S}^d$  were replaced by  $\mathbb{R}^{d+1}$ . Nevertheless, the high frequency behavior of  $\mathcal{L}_\beta$  is the same as the Laplace-Beltrami operator  $(\Delta_d)^\beta$  as will be clear in the proof of Lemma 4.1. As in the diffusion limit, we also observe slightly stronger convergence results with the probabilistic techniques, in particular pointwise convergence in  $(t, x, k)$  for bounded and continuous initial conditions. Note as well that the solutions to (9) are  $\mathcal{C}^\infty$  according to Theorem 3.1.

## 4 Proof of Theorem 3.1

The proof is divided into several steps. We first derive some important results on the operator  $\mathcal{L}$ . In a second time, we regularize the kernel  $F$  in order to show the existence and uniqueness of weak solutions along with some energy estimates. We finally obtain the  $C^\infty$  regularity using Theorem 3.1, first with respect to the  $(t, x)$  variables, and then with respect to the momentum  $k$ .

### 4.1 Step 0: Properties of the operator $\mathcal{L}$

**Lemma 4.1** *The operator  $\mathcal{L}$  defined on  $C^\infty(\mathbb{S}^d)$  satisfies the following properties:*

- (i) For any  $f \in C^\infty(\mathbb{S}^d)$ ,  $\mathcal{L}f \in L^\infty(\mathbb{S}^d)$ .
- (ii)  $\mathcal{L}$  can be extended to a unique non-positive self-adjoint operator  $\bar{\mathcal{L}}$  with domain  $\mathcal{D}(\bar{\mathcal{L}}) = \{\varphi \in H^\beta(\mathbb{S}^d), \bar{\mathcal{L}}\varphi \in L^2(\mathbb{S}^d)\}$ . Moreover,  $-\bar{\mathcal{L}}$  is associated to a quadratic form  $\mathcal{Q} : H^\beta(\mathbb{S}^d) \times H^\beta(\mathbb{S}^d) \rightarrow \mathbb{R}_+$  given by

$$\mathcal{Q}(f, g) = \frac{1}{2} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} F(k \cdot p)(f(k) - f(p))(g(k) - g(p)) d\sigma(k) d\sigma(p),$$

which satisfies the estimate, for some  $C > 0$ ,

$$C\|f\|_{H^\beta(\mathbb{S}^d)}^2 \leq \mathcal{Q}(f, f) + \|f\|_{L^2(\mathbb{S}^d)}^2 \leq C^{-1}\|f\|_{H^\beta(\mathbb{S}^d)}^2, \quad \forall f \in H^\beta(\mathbb{S}^d). \quad (10)$$

Also,  $\mathcal{L}$  can be extended to a bounded operator (still denoted by  $\mathcal{L}$  for simplicity) from  $H^\beta(\mathbb{S}^d)$  to its dual  $(H^\beta(\mathbb{S}^d))^*$ .

- (iii) Let  $(\cdot)_-$  denote the negative part of a function. Assume there exists  $b \in (0, a_1)$  such that the kernel  $F$  verifies

$$(F(s) - a_1|1 - s|^{-\beta-d/2})_- \leq b|1 - s|^{-\beta-d/2}, \quad a.e. \text{ on } (-1, 1).$$

Then, there exists a constant  $C > 0$  such that, for  $e_0 = \sigma(\mathbb{S}^d)^{-1/2}$ ,

$$C\|f - \langle f, e_0 \rangle\|_{H^\beta(\mathbb{S}^d)}^2 \leq \mathcal{Q}(f, f). \quad (11)$$

- (iv) For any  $f \in H^\beta(\mathbb{S}^d)$ , there exists  $h \in L^2(\mathbb{S}^d)$  such that

$$(-\Delta_d)^{\beta/2} h = \mathcal{L}f, \quad \text{with} \quad \|h\|_{L^2(\mathbb{S}^d)} \leq C\|f\|_{H^\beta(\mathbb{S}^d)}.$$

*Proof.* (i) Let  $f \in C^\infty(\mathbb{S}^d)$ . We have, for all  $k \in \mathbb{S}^d$ ,

$$\mathcal{L}f(k) = \lim_{\eta \rightarrow 0} \int_{|k-p| > \sqrt{2}\eta} F(k \cdot p)(f(p) - f(k)) d\sigma(p).$$

We split  $\mathcal{L}$  into  $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$  according to (4), and only treat  $\mathcal{L}_2$  since the term  $\mathcal{L}_1$  is straightforward. We may assume without loss of generality that  $k = e_{d+1} = (0, \dots, 0, 1) \in \mathbb{R}^{d+1}$ , and write, when  $d \geq 2$ ,

$$p = (\sqrt{1 - s^2}u, s),$$

with  $s \in [-1, 1]$ , and  $u \in \mathbb{S}^{d-1}$ . Then,

$$\mathcal{L}_2 f(k) = \lim_{\eta \rightarrow 0} \int_{\mathbb{S}^{d-1}} \int_{-1}^{1-\eta} F_2(s) (f(\sqrt{1-s^2}u + sk) - f(k)) (1-s^2)^{\frac{d-2}{2}} d\sigma(u) ds. \quad (12)$$

Recasting  $f$  as

$$f(k) = \varphi(0, \dots, 0, 1), \quad f(\sqrt{1-s^2}u + sk) = \varphi(\sqrt{1-s^2}u_1, \dots, \sqrt{1-s^2}u_d, s),$$

we have

$$\begin{aligned} f(\sqrt{1-s^2}u + sk) - f(k) &= (s-1) \partial_{x_{d+1}} \varphi(0, \dots, 0, 1) + \sqrt{1-s^2}u \cdot \nabla_d \varphi(0, \dots, 0, 1) \\ &\quad + \mathcal{O}(|s-1|), \end{aligned}$$

where  $\partial_{x_{d+1}}$  denotes partial derivation with respect to the  $(d+1)$ -th variable and  $\nabla_d$  the gradient with respect to the  $d$  first variables. Since  $\int_{\mathbb{S}^{d-1}} u d\sigma(u) = 0$ , it follows from the equation above and (4) that

$$|\mathcal{L}_2 f(k)| \leq C \int_{-1}^1 \frac{(1-s)(1-s^2)^{\frac{d-2}{2}}}{(1-s)^{\beta+\frac{d}{2}}} ds \leq C \int_{-1}^1 \frac{ds}{(1-s)^\beta} \leq C, \quad (13)$$

since  $\beta \in (0, 1)$ . This proves item (i) for  $d \geq 2$ . The case  $d = 1$  follows analogously after a simple adaptation.

(ii) Let  $f$  and  $g$  be in  $\mathcal{C}^\infty(\mathbb{S}^d)$ . Then

$$\begin{aligned} \langle \mathcal{L}f, g \rangle &= \text{p.v.} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} F(k \cdot p) (f(p) - f(k)) g(k) d\sigma(p) d\sigma(k) \\ &= \frac{1}{2} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} F(k \cdot p) (f(p) - f(k)) (g(k) - g(p)) d\sigma(p) d\sigma(k) \\ &= -\mathcal{Q}(f, g). \end{aligned} \quad (14)$$

Using (4) and a Cauchy-Schwarz inequality, we find

$$|\langle \mathcal{L}_2 f, g \rangle|^2 \leq C \left( \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \frac{|f(k) - f(p)|^2}{|k-p|^{2\beta+d}} d\sigma(k) d\sigma(p) \right) \left( \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \frac{|g(k) - g(p)|^2}{|k-p|^{2\beta+d}} d\sigma(k) d\sigma(p) \right).$$

Introducing the following operator, for  $\beta \in (0, 1)$ ,

$$\mathcal{R}_\beta f(k) = \text{p.v.} \int_{\mathbb{S}^d} \frac{f(k) - f(p)}{|k-p|^{2\beta+d}} d\sigma(p), \quad k \in \mathbb{S}^d,$$

we have

$$\int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \frac{|f(k) - f(p)|^2}{|k-p|^{2\beta+d}} d\sigma(k) d\sigma(p) = 2 \langle \mathcal{R}_\beta f, f \rangle.$$

It is proved in [27] (with a slight adaptation of the constants), that the Fourier multipliers  $R_n$  associated with  $\mathcal{R}_\beta$  are

$$R_n = \frac{2^{2\beta} \pi^{\frac{d}{2}} \Gamma(\beta)}{\Gamma(\frac{d}{2} + \beta)} \left( \frac{\Gamma(n + \frac{d+2\beta}{2})}{\Gamma(n + \frac{d-2\beta}{2})} - \frac{\Gamma(\frac{d+2\beta}{2})}{\Gamma(\frac{d-2\beta}{2})} \right),$$

where  $\Gamma$  is the gamma function. When  $d = 1$  and  $\beta = 1/2$ , we have, by convention,  $\Gamma(\frac{d-2\beta}{2}) = \Gamma(0) = \infty$ . The fact that  $\Gamma(n + \alpha) \sim \Gamma(n)n^\alpha$  as  $n \rightarrow \infty$  for  $\alpha \in \mathbb{R}$ , shows that  $R_n$  behaves like  $n^{2\beta}$  for large  $n$ , which is the same asymptotics as the eigenvalues of  $(-\Delta_d)^\beta$ . We can therefore write

$$\langle \mathcal{R}_\beta f, f \rangle \leq C \|f\|_{H^\beta(\mathbb{S}^d)}^2,$$

so that

$$|\langle \mathcal{L}f, g \rangle| \leq C \|f\|_{H^\beta(\mathbb{S}^d)} \|g\|_{H^\beta(\mathbb{S}^d)}. \quad (15)$$

By density of  $\mathcal{C}^\infty(\mathbb{S}^d)$  in  $H^\beta(\mathbb{S}^d)$ , Theorem X.23 of [26] then shows that the quadratic form  $\mathcal{Q}$  can be extended to a quadratic form (still denoted by  $\mathcal{Q}$ ) with domain  $\mathcal{D}(\mathcal{Q}) = H^\beta(\mathbb{S}^d)$ , associated to a unique self-adjoint operator  $\bar{\mathcal{L}}$ , with the domain given in (ii), such that

$$\mathcal{Q}(f, g) = -\langle \bar{\mathcal{L}}f, g \rangle, \quad \forall f \in \mathcal{D}(\bar{\mathcal{L}}), \quad \forall g \in H^\beta(\mathbb{S}^d).$$

The upper bound in (10) follows from (15). For the lower bound, we simply remark that there are positive constants  $C_1$  and  $C_2$  such that

$$\langle \mathcal{R}_\beta f, f \rangle \leq C_1 \langle \mathcal{L}_2 f, f \rangle + C_2 \|f\|_{L^2(\mathbb{S}^d)}^2$$

and use the asymptotic behavior of  $R_n$ . Note that we can also conclude from (15) that  $\mathcal{L}$  can be extended to a bounded operator (still denoted by  $\mathcal{L}$ ) from  $H^\beta(\mathbb{S}^d)$  to its dual  $(H^\beta(\mathbb{S}^d))^*$ .

(iii) For  $f \in H^\beta(\mathbb{S}^d)$ , we have

$$\begin{aligned} 2\mathcal{Q}(f, f) &= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \left( \frac{a_1}{|k-p|^{2\beta+d}} + F(k \cdot p) - \frac{a_1}{|k-p|^{2\beta+d}} \right) (f(k) - f(p))^2 d\sigma(k) \sigma(p) \\ &\geq \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \left( \frac{a_1}{|k-p|^{2\beta+d}} - \left( F(k \cdot p) - \frac{a_1}{|k-p|^{2\beta+d}} \right) \right) (f(k) - f(p))^2 d\sigma(k) \sigma(p) \\ &\geq \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \frac{a_1 - b}{|k-p|^{2\beta+d}} (f(k) - f(p))^2 d\sigma(k) \sigma(p) \\ &\geq 2(a_1 - b) \langle \mathcal{R}_\beta f, f \rangle. \end{aligned}$$

The conclusion follows once more from the asymptotics of the multipliers  $R_n$ , and the fact that  $R_0 = 0$ .

(iv) Let  $f \in \mathcal{C}^\infty(\mathbb{S}^d)$ . Let us remark first that if  $(-\Delta_d)^\gamma : H^{2\gamma}(\mathbb{S}^d) \rightarrow L^2(\mathbb{S}^d)$ , with  $\gamma > 0$ , then

$$\ker(-\Delta_d)^\gamma = \mathbb{C}.$$

Hence,  $(-\Delta_d)^{-\gamma}$ , the inverse of  $(-\Delta_d)^\gamma$ , is defined on the set of functions  $\varphi \in L^2(\mathbb{S}^d)$  with vanishing integral on the sphere, and  $\int_{\mathbb{S}^d} (-\Delta_d)^{-\gamma} \varphi(k) d\sigma(k) = 0$ . Then, since  $\int_{\mathbb{S}^d} \mathcal{L}f(k) d\sigma(k) = 0$ , we can write

$$\mathcal{L}f = (-\Delta_d)^{\frac{\beta}{2}} (-\Delta_d)^{-\frac{\beta}{2}} \mathcal{L}f.$$

It is then direct to see that the operator  $(-\Delta_d)^{-\frac{\beta}{2}}\mathcal{L}$  can be extended to a bounded operator  $A$  from  $H^\beta(\mathbb{S}^d)$  to  $L^2(\mathbb{S}^d)$ . Denote indeed by  $\langle g \rangle$  the integral of  $g \in \mathcal{C}^\infty(\mathbb{S}^d)$  on  $\mathbb{S}^d$ , so that

$$\langle (-\Delta_d)^{-\frac{\beta}{2}}\mathcal{L}f, g \rangle = \langle (-\Delta_d)^{-\frac{\beta}{2}}\mathcal{L}f, g - \langle g \rangle \rangle = \langle \mathcal{L}f, (-\Delta_d)^{-\frac{\beta}{2}}(g - \langle g \rangle) \rangle,$$

we then deduce from (15) that

$$|\langle \mathcal{L}f, (-\Delta_d)^{-\frac{\beta}{2}}(g - \langle g \rangle) \rangle| \leq C \|f\|_{H^\beta(\mathbb{S}^d)} \|(-\Delta_d)^{-\frac{\beta}{2}}(g - \langle g \rangle)\|_{H^\beta(\mathbb{S}^d)} \leq C \|f\|_{H^\beta(\mathbb{S}^d)} \|g\|_{L^2(\mathbb{S}^d)}.$$

It finally suffices to set  $h = Af$  to conclude the proof.  $\square$

## 4.2 Step 1: standard estimates for the Cauchy problem

**Proposition 4.2** *The radiative transfer equation (1) admits a unique weak solution  $f$  that satisfies  $(-\Delta_d)^{\frac{\beta}{2}}f \in L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  and the following estimate: for any  $T > 0$ ,*

$$\|f\|_{L^\infty((0,T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))}^2 + 2 \int_0^T \mathcal{Q}(f, f) dt \leq \|f_0\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}^2, \quad (16)$$

where the quadratic form  $\mathcal{Q}$  is defined in Lemma 4.1. Moreover, if for any  $\mathcal{C}^\infty$  function  $\chi$  with compact support in  $(0, \infty)$ , we have  $\partial_{t,x}^\gamma f_\chi \in L^2(\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d)$ , where  $\gamma > 0$  and  $f_\chi$  is extension by zero of  $\chi f$  to all  $t$  in  $\mathbb{R}$ , then

$$(-\Delta_d)^{\frac{\beta}{2}} \partial_{t,x}^\gamma f_\chi \in L^2(\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d). \quad (17)$$

*Proof.* We consider a slightly more general version of (1) that will be helpful for the adjoint problem and for (17): we suppose there is a source term in the transport equation so that  $f$  satisfies

$$\partial_t f + k \cdot \nabla_x f = \mathcal{L}f + S, \quad S \in L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d). \quad (18)$$

We then proceed by regularization: let  $f_n^0 \in \mathcal{C}^\infty(\mathbb{R}^{d+1} \times \mathbb{S}^d)$  with compact support in  $x$ ,  $S_n \in \mathcal{C}^\infty([0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  with compact support in  $(t, x)$  such that, as  $n \rightarrow \infty$ ,

$$\begin{aligned} f_n^0 &\rightarrow f^0, & \text{strongly in } & L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d) \\ S_n &\rightarrow S, & \text{strongly in } & L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d). \end{aligned}$$

The regularization of  $F$  is done as follows: consider decomposition (4). The functions  $F_1$  and  $a_2$  are regularized into  $\mathcal{C}^\infty([-1, 1])$  functions with usual mollifiers as  $F_{1,n}$  and  $a_{2,n}$  with

$$F_{1,n} \rightarrow F_1, \quad \text{strongly in } L^1((-1, 1)) \quad (19)$$

$$a_{2,n} \rightarrow a_2, \quad \text{a.e. on } (\delta, 1). \quad (20)$$

The function  $F_2$  is regularized as

$$F_{2,n}(s) = \frac{a_1(1 + a_{2,n}(s))(1 - \chi(s))}{|1 - s + n^{-1}|^{\beta + \frac{d}{2}}},$$

with  $1 + a_{2,n}(s) \geq 0$  in  $[\delta, 1]$  since  $a_{2,n}$  is such that  $\|a_{2,n}\|_{L^\infty((\delta,1))} \leq \|a_2\|_{L^\infty((\delta,1))} < 1$ . It is then direct to obtain the existence of a smooth solution  $f_n \in \mathcal{C}^\infty([0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  with compact support in  $x$  (using for instance Duhamel expansions and the fact that the collision kernel is bounded). Multiplying the regularized transport equation by  $f_n$  and integrating in all variables leads to, for any  $T \in (0, \infty)$ ,

$$\begin{aligned} \|f_n\|_{L^\infty((0,T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))}^2 + \sum_{i=1}^2 \|h_{i,n}\|_{L^2((0,T) \times \mathbb{R}^{d+1} \times \mathbb{S}^d \times \mathbb{S}^d)}^2 & \quad (21) \\ \leq \|f_n^0\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}^2 + 2\|S_n\|_{L^2((0,T) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)} \|f_n\|_{L^2((0,T) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)} \end{aligned}$$

where

$$h_{i,n}(t, x, k, p) = \sqrt{F_{i,n}(k \cdot p)} (f_n(t, x, k) - f_n(t, x, p)). \quad (22)$$

Since by construction,

$$\|f_n^0\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)} \leq \|f^0\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}, \quad \|S_n\|_{L^2((0,\infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)} \leq \|S\|_{L^2((0,\infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)},$$

we can deduce uniform bounds from (21), and standard compactness arguments show that one can extract subsequences such that, for  $i = 1, 2$ :

$$f_n \rightarrow f \quad \text{weak-}\star \text{ in } \quad L^\infty((0, \infty), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)) \quad (23)$$

$$h_{i,n} \rightarrow h_i \quad \text{weakly in } \quad L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d \times \mathbb{S}^d). \quad (24)$$

It remains now to identify the limits. We focus only on the most interesting terms  $f$  and  $h_2$  and leave the details for  $h_1$  to the reader. We start with  $h_{2,n}$ . Let  $\varphi \in L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d \times \mathbb{S}^d)$  and denote by  $(\cdot, \cdot)$  the corresponding inner product. Functions of the form  $\varphi_1(t) \otimes \varphi_2(x) \otimes \varphi_3(k) \otimes \varphi_4(p)$ , with  $\varphi_i$  infinitely differentiable with compact support, are dense in  $L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d \times \mathbb{S}^d)$ , see [29], Chapter 39, so it is enough to consider a test function  $\varphi$  of the latter tensor form. We then have

$$(h_{2,n}, \varphi) = \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} f_n(t, x, k) \varphi_1(t) \varphi_2(x) \psi_n(k, p) dt dx d\sigma(k) d\sigma(p),$$

where

$$\psi_n(k, p) = \sqrt{F_{2,n}(k \cdot p)} (\varphi_3(k) \varphi_4(p) - \varphi_3(p) \varphi_4(k)).$$

Owing the weak convergence (23), we need to show the strong convergence of  $\psi_n$  in  $L^2$  in order to identify the limit. Since  $\psi_n$  can be written as

$$\psi_n(k, p) = \sqrt{F_{2,n}(k \cdot p)} (\varphi_3(k) \varphi_4(k) - \varphi_3(p) \varphi_4(p) - (\varphi_3(k) + \varphi_4(p)) (\varphi_4(k) - \varphi_4(p))),$$

it is enough to show the strong convergence in  $L^2(\mathbb{S}^d \times \mathbb{S}^d)$  of functions of the form

$$\widetilde{\psi}_n(k, p) = \sqrt{F_{2,n}(k \cdot p)} (\varphi(k) - \varphi(p))$$

where  $\varphi \in \mathcal{C}^\infty(\mathbb{S}^d)$ . For this, we remark first that by construction, using (20), as  $n \rightarrow \infty$ ,

$$F_{2,n}(s) \rightarrow F_2(s) \quad \text{a.e. on } (\delta, 1). \quad (25)$$

Introduce then

$$I_n := \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \left( \sqrt{F_{2,n}(k \cdot p)} - \sqrt{F_2(k \cdot p)} \right)^2 (\varphi(k) - \varphi(p))^2 d\sigma(k) d\sigma(p).$$

Since

$$|\varphi(k) - \varphi(p)| \leq C|k - p|,$$

and  $F_{2,n}, F_2$  are supported on  $[\delta, 1]$ , we have

$$I_n \leq C \int_{\delta}^1 \left( \frac{\sqrt{1 + a_2(s)}}{|1 - s|^{\frac{\beta}{2} + \frac{d}{4}}} - \frac{\sqrt{1 + a_{2,n}(s)}}{|1 - s + n^{-1}|^{\frac{\beta}{2} + \frac{d}{4}}} \right)^2 (1 - s)(1 - s^2)^{\frac{d-2}{2}} ds.$$

The integrand being uniformly bounded by the function  $C(1 - t)^{-\beta} \in L^1((-1, 1))$  since  $\beta \in (0, 1)$ , (25) together with dominated convergence show that  $I_n \rightarrow 0$ . This implies that

$$(h_{2,n}, \varphi) \rightarrow \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} f(t, x, k) \varphi_1(t) \varphi_2(x) \psi(k, p) dt dx d\sigma(k) d\sigma(p),$$

with obvious notation for  $\psi$ , which yields

$$h_2(t, x, k, p) = \sqrt{F_2(k \cdot p)} (f(t, x, k) - f(t, x, p)).$$

From (21), (23) and (24), we then deduce (16) when  $S = 0$ . It only remains to pass to the limit in the weak formulation. The only term requiring some attention is the one involving the collision kernel, for which we need to show that terms of the form

$$\text{p.v.} \int_{\mathbb{S}^d} (F_{1,n} + F_{2,n} - F_1 - F_2)(k \cdot p) (\varphi(p) - \varphi(k)) d\sigma(p)$$

converge to zero strongly in  $L^2(\mathbb{S}^d)$  for  $\varphi$  smooth. This is done using (25) and the decomposition (12) in order to get a majorizing function for dominated convergence. Summarizing, we have therefore obtained the existence of a weak solution  $f$  satisfying estimate (16). The latter, together with (11), yields  $(-\Delta_d)^{\frac{\beta}{2}} f \in L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ .

Let us prove now the uniqueness. We proceed as usual with the adjoint problem, and need to show that for  $f$  a weak solution with a vanishing initial condition, the following equality, for all smooth  $\varphi$  with compact support in  $x$  such that  $\varphi = 0$  for  $t \geq T$ ,  $T$  arbitrary,

$$\int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f(t, x, k) (\partial_t + k \cdot \nabla_x + \mathcal{L}) \varphi(t, x, k) dt dx d\sigma(k) = 0,$$

implies  $f = 0$ . Note that for technical reasons, the latter condition can be recast as

$$\lim_{n \rightarrow \infty} \int_{n^{-1}}^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f(t, x, k) (\partial_t + k \cdot \nabla_x + \mathcal{L}) \varphi(t, x, k) dt dx d\sigma(k) = 0. \quad (26)$$

Let  $v \in L^\infty((0, T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))$  be a weak solution to the problem

$$(\partial_t + k \cdot \nabla_x - \mathcal{L})v(t, x, k) = -f(T - t, x, -k), \quad v(0, x, k) = 0,$$



and define  $u(t, x, k) = v(T - t, x, -k)$ . We then verify that  $u$  is a weak solution to the adjoint problem

$$(\partial_t + k \cdot \nabla_x + \mathcal{L})u(t, x, k) = f(t, x, k), \quad u(T, x, k) = 0,$$

that is, for all  $\varphi$  smooth with  $\varphi(0, x, k) = 0$ ,

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} u(t, x, k) (-\partial_t - k \cdot \nabla_x + \mathcal{L})\varphi(t, x, k) dt dx d\sigma(k) \\ &= \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f(t, x, k) \varphi(t, x, k) dt dx d\sigma(k). \end{aligned} \quad (27)$$

If  $u$  were smooth, we would plug  $\varphi = u$  in (26), which would lead to  $f = 0$ . We only know that  $u \in L^\infty((0, T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))$  and thus need to regularize. Let  $u_n$  be a regularized version of  $u$  defined by, for  $t \in [n^{-1}, T]$ ,

$$u_n(t, x, k) = \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} \varphi_n(t-s) \phi_n(x-y) \psi_n(k \cdot p) u(s, y, p) ds dy d\sigma(p),$$

where  $\varphi_n$  is such that  $\varphi_n(u) = 0$  for  $|u| \geq n^{-1}$ . We have consequently,

$$\begin{aligned} & (\partial_t + k \cdot \nabla_x + \mathcal{L})u_n(t, x, k) \\ &= \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} u(s, y, p) (-\partial_s - k \cdot \nabla_y + \mathcal{L})\varphi_n(t-s) \phi_n(x-y) \psi_n(k \cdot p) ds dy d\sigma(p). \end{aligned}$$

Since  $\varphi_n(t) = 0$  for  $t \in [n^{-1}, T]$ , we can use the fact that  $u$  is a weak solution, and therefore (27), to write, for all  $t \in [n^{-1}, T]$ ,

$$(\partial_t + k \cdot \nabla_x + \mathcal{L})u_n(t, x, k) = f_n(t, x, k),$$

where

$$f_n(t, x, k) = \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f(s, y, p) \varphi_n(t-s) \phi_n(x-y) \psi_n(k \cdot p) ds dy d\sigma(p).$$

Assuming the regularization is such that

$$f_n \rightarrow f, \quad \text{strongly in } L^2((0, T) \times \mathbb{R}^{d+1} \times \mathbb{S}^d),$$

it is then straightforward to conclude from (26) that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{n^{-1}}^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f(t, x, k) (\partial_t + k \cdot \nabla_x + \mathcal{L})u_n(t, x, k) dt dx d\sigma(k) \\ &= \lim_{n \rightarrow \infty} \int_{n^{-1}}^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} f(t, x, k) f_n(t, x, k) dt dx d\sigma(k) \\ &= \int_0^T \int_{\mathbb{R}^{d+1}} \int_{\mathbb{S}^d} (f(t, x, k))^2 dt dx d\sigma(k) = 0, \end{aligned}$$

which proves the uniqueness.

Regarding (17), suppose  $\partial_{t,x}^\gamma f_\chi \in L^2(\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d)$ , and consider the regularized  $f_n$  obtained previously. We remark that  $u_n = \partial_{t,x}^\gamma f_{n,\chi}$  ( $f_{n,\chi}$  the extension by zero of  $\chi f_n$  to all  $t \in \mathbb{R}$ ) satisfies (18) on  $\mathbb{R}_t \times \mathbb{R}_x^{d+1} \times \mathbb{S}^d$  with  $f = u_n$  and  $S = S_n = -\partial_{t,x}^\gamma f_{n,\chi}$ . Note also that  $\|u_n(t)\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)} \rightarrow 0$  as  $|t| \rightarrow \infty$ . Multiplying then (18) by  $u_n$  and integrating over all variables leads to

$$\frac{1}{2} \sum_{i=1}^2 \|h_{i,n}\|_{L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d \times \mathbb{S}^d)}^2 \leq \|\partial_{t,x}^\gamma f_{n,\chi'}\|_{L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)} \|u_n\|_{L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)} \quad (28)$$

where  $h_{i,n}$  is as in (22) with  $f_n$  replaced by  $u_n$ . We showed above that  $f_n$  converges weakly to the unique weak solution  $f$ . It is then not hard to see that the fact that  $\partial_{t,x}^\gamma f_\chi$  belongs to  $L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  for any smooth  $\chi$  with compact support in  $(0, \infty)$  implies that  $S_n$  and  $u_n$  are uniformly bounded in  $L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . Following along the same lines as in the proof of (16), we can finally pass to the limit in (28) and obtain that

$$\int_{\mathbb{R}} \mathcal{Q}(u, u) dt \leq \|\partial_{t,x}^\gamma f_{\chi'}\|_{L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)} \|u\|_{L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)}, \quad u = \partial_{t,x}^\gamma f_\chi,$$

which yields (17) thanks to (11).  $\square$

### 4.3 Step 2: Regularity in $(t, x)$

**Proposition 4.3** *Let  $f$  be the weak solution to (1) and let  $\gamma = \frac{\beta}{2+3\beta}$ . Then, for any  $n \geq 0$ , any  $m \geq 0$ , and any  $\mathcal{C}^\infty$  function  $\chi$  with compact support in  $(0, \infty)$ , we have*

$$\partial_t^{n\gamma} \partial_{x_j}^{m\gamma} f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1}, H^\beta(\mathbb{S}^d)), \quad j = 1, \dots, d+1,$$

where  $f_\chi$  is the extension by zero of the function  $\chi f$  to all  $t \in \mathbb{R}$ .

*Proof.* We know from Proposition 4.2 that the result is true for  $n = m = 0$ , and prove the proposition by a double induction on  $n$  and  $m$ . Let us show first that  $\partial_t^{n\gamma} f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . By construction, the function  $f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  verifies in the distributional sense

$$\partial_t f_\chi + k \cdot \nabla_x f_\chi = \mathcal{L} f_\chi + S, \quad S = -f_{\chi'}.$$

We start with the case  $n = 1$ . We know from Proposition 4.2 that  $(-\Delta_d)^{\frac{\beta}{2}} f \in L^2((0, \infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ , so that  $(-\Delta_d)^{\frac{\beta}{2}} f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . We have as well  $S \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . In order to apply Theorem 1.1 and the hypoelliptic estimates, we use item (iv) of Proposition 4.1 to write  $\mathcal{L} f_\chi = (-\Delta_d)^{\frac{\beta}{2}} h$ , for some  $h \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . It then follows from Theorem 1.1 (with  $\theta \equiv \beta$  and  $\beta \equiv \beta/2$ ) that  $\partial_t^\gamma f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ , which proves the case  $n = 1$ .

Suppose now the result holds true at the step  $n$ ,  $n \geq 1$ , so that for any  $\chi$  with the properties stated in the proposition, we have  $u := \partial_t^{n\gamma} f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . Result (17) of Proposition 4.2 then yields

$$(-\Delta_d)^{\frac{\beta}{2}} u \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d).$$

In order to conclude, we remark that the function  $u$  satisfies in the distributional sense

$$\partial_t u + k \cdot \nabla_x u = \mathcal{L} u + S, \quad S = -\partial_{t,x}^{n\gamma} f_{\chi'}.$$

Since  $\chi' \in \mathcal{C}^\infty$  with compact support in  $(0, \infty)$ , we have by the induction hypothesis that  $S \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . Applying once more Theorem 1.1, we find  $\partial_t^\gamma u \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ , which proves the result for the step  $n+1$  and therefore for all  $n$ . Following along the same lines shows that  $\partial_{x_j}^{m\gamma} f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  for all  $m$ , which completes the initialization step of the double induction.

The evolution step from  $(n, m)$  to  $(n+1, m)$ , to  $(n, m+1)$  and to  $(n+1, m+1)$  is done in a similar fashion as above, by defining first  $u$  by  $u := \partial_t^{n\gamma} \partial_{x_j}^{m\gamma} f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$  and then by  $u := \partial_t^{(n+1)\gamma} \partial_{x_j}^{m\gamma} f_\chi \in L^2(\mathbb{R} \times \mathbb{R}^{d+1} \times \mathbb{S}^d)$ . This ends the proof.  $\square$

#### 4.4 Step 3: Regularity in $k$ and conclusion

Take  $t_0 > 0$  and  $T > 0$  arbitrary with  $t_0 < T$ , and let  $\chi \in \mathcal{C}^\infty(\mathbb{R})$  such that  $\chi(t) = 1$  for  $t \in [t_0, T]$ , and  $\chi(t) = 0$  for  $t \leq t_0/2$  and  $t \geq 2T$ . It follows from Proposition 4.3 that, for all  $n \geq 0$  and  $m \geq 0$ ,

$$\partial_t^m \partial_{x_j}^n f \in L^2((t_0, T) \times \mathbb{R}^{d+1}, H^\beta(\mathbb{S}^d)), \quad j = 1, \dots, d+1.$$

Since  $t_0$  and  $T$  are arbitrary, standard Sobolev embeddings then yield

$$f \in \mathcal{C}_{t,x}^\infty((0, \infty) \times \mathbb{R}^{d+1}, H^\beta(\mathbb{S}^d)).$$

It thus remains to show the regularity in the  $k$  variable. Define for this the operator  $A := -\tilde{\mathcal{L}} + \mathbb{I}$ . It is direct to see that  $A$  is a sectorial operator that generates a holomorphic semigroup  $S(t)$ . For  $\theta \in (0, \pi/2]$ , let indeed  $S_\theta = \{z \in \mathbb{C}, |\arg(z)| \leq \pi/2 - \theta, \Re(z) \geq 0\}$ . The spectrum of  $A$ , denoted by  $\sigma(A)$ , satisfies  $\sigma(A) \subset [1, \infty)$  and therefore  $\sigma(A) \subset S_\theta$ , for all  $\theta \in (0, \pi/2]$ . Hence, for  $z \in \mathbb{C} \setminus S_{\theta_1}$  with  $0 < \theta_1 < \theta$ , we have,

$$\|(z + A)^{-1}\|_{L^2(\mathbb{S}^d)} = \frac{1}{\text{dist}(z, \sigma(A))} \leq \frac{1}{\text{dist}(z, S_{\theta_1})},$$

and it suffices to apply Theorem X.25 of [26] (note that  $A$  is closed since it is self-adjoint). It then follows from [26], Corollary 2, page 252, that, for any  $\varphi \in L^2(\mathbb{S}^d)$ ,  $S(t)\varphi \in D(A^p)$ , for all  $t > 0$  and  $p \in \mathbb{N}^*$ , with the estimate

$$\|A^p S(t)\varphi\|_{L^2(\mathbb{S}^d)} \leq C t^{-p} \|\varphi\|_{L^2(\mathbb{S}^d)}. \quad (29)$$

By interpolation, the latter estimate can be extended to all  $p$  real and positive. Let then  $u := \partial_t^m \partial_{x_j}^n f$ ,  $j = 1, \dots, d+1$ ,  $n$  and  $m$  arbitrary, let  $\chi$  defined as before, and let  $v := -\chi k \cdot \nabla_x u - u(\chi' - 1)$ . Since  $f$  is smooth in  $(t, x)$ ,  $v$  is defined for all  $(t, x)$  and a.e. in  $\mathbb{S}^d$ . This allows us to use the following representation formula for  $u\chi$ ,

$$u(t)\chi(t) = \int_0^t S(t-s)v(s)ds.$$

We show now that for all  $(m, n, j)$ ,  $A^{p/2}u \in L^\infty((0, T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))$ , for all  $p \in \mathbb{N}^*$ . We proceed by induction, and start with the initial step  $p = 1$ . According to (10), we have, for all  $(t, x)$ ,

$$\langle Au, u \rangle = \langle A^{1/2}u, A^{1/2}u \rangle \leq C \|u\|_{H^\beta(\mathbb{S}^d)}^2,$$

which proves the claim together with Proposition 4.3. Suppose now that for all  $(m, n, j)$ ,  $A^{p/2}u \in L^\infty((0, T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))$ . We have

$$A^{(p+1)/2}u(t)\chi(t) = \int_0^t A^{1/2}S(t-s)A^{p/2}v(s)ds.$$

Above, we used the fact that  $A^{(p+1)/2}$  is closed, and that  $S$  and  $A^{p/2}$  commute. Using then (29), we find, for  $0 < t \leq T$ ,

$$\|A^{(p+1)/2}u(t)\chi(t)\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)} \leq C\|A^{p/2}v\|_{L^\infty((0, T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))} \int_0^t (t-s)^{-1/2}ds,$$

which is finite for all  $(m, n, j)$  according to the induction hypothesis and proves the claim.

We conclude the proof of the proposition by showing that  $(-\Delta_d)^{p\beta/2}u \in L^\infty((0, T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d))$  for all  $p \in \mathbb{N}^*$ . We write for this, using (10), for all  $p \geq 1$ ,

$$C\|A^p u\|_{H^\beta(\mathbb{S}^d)}^2 \leq \langle AA^p u, A^p u \rangle = \langle A^{2p+1}u, u \rangle = \|A^{p+1/2}u\|_{L^2(\mathbb{S}^d)}^2. \quad (30)$$

Using the Funk-Hekke formula [6], Chapter 2, page 36, it is not difficult to see that  $A$  and  $(-\Delta_d)^{\frac{\beta}{2}}$  commute, so that

$$\begin{aligned} \|A^p u\|_{H^\beta(\mathbb{S}^d)}^2 &= \|A^p u\|_{L^2(\mathbb{S}^d)}^2 + \|(-\Delta_d)^{\beta/2}A^p u\|_{L^2(\mathbb{S}^d)}^2 \\ &= \|A^p u\|_{L^2(\mathbb{S}^d)}^2 + \|A^p(-\Delta_d)^{\beta/2}u\|_{L^2(\mathbb{S}^d)}^2. \end{aligned}$$

With (30), this finally yields

$$\|A^p(-\Delta_d)^{\beta/2}u\|_{L^2(\mathbb{S}^d)} \leq C\|A^{p+1/2}u\|_{L^2(\mathbb{S}^d)},$$

and it suffices to iterate to conclude that, for any  $0 < t_0 < T$ , any  $(m, n, j, p)$ ,

$$(-\Delta_d)^{p\beta/2}\partial_t^m \partial_{x_j}^n f \in L^\infty((t_0, T), L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)).$$

The proof is ended with classical Sobolev embeddings.

## 5 Proof of Theorem 3.2

The proof is based on standard convergence arguments for stochastic processes. The main ingredient is the spectral gap estimate (iii) of Lemma 4.1.

**Step 1: spectral gap and invariant measure.** Let us first verify that the hypothesis on the kernel  $F$  in item (iii) is verified. According to (3), for all  $b > 0$ , there exists  $\delta_0$  such that

$$|F(s) - a_1|1 - s|^{-\beta-d/2}| \leq b|1 - s|^{-\beta-d/2}, \quad \forall s \in [\delta_0, 1).$$

It remains to treat the case  $s \in (-1, \delta_0)$ . It is assumed here that there exists  $\eta$  such that  $F(x) \geq \eta > 0$  a.e. on  $(-1, 1)$ . Suppose  $F(s)(1-s)^{\beta+d/2} - a_1 \leq 0$  for a non empty subset  $I$  of  $(-1, \delta_0)$ , otherwise we are done. Choose then  $b < a_1$  such that the associated  $\delta_0$  yields  $a_1 - \eta(1 - \delta_0)^{\beta+d/2} > 0$ . Then, for all  $s \in I$ ,

$$a_1 - F(s)(1-s)^{\beta+d/2} \leq a_1 - \eta(1 - \delta_0)^{\beta+d/2},$$

and the hypothesis is verified. The next lemma is a direct consequence of Lemma 4.1 item (iii):

**Lemma 5.1** *We have the following properties for the operator  $\bar{\mathcal{L}}$  and the Markov process  $\mathbf{M}$ .*

1. *0 is a simple eigenvalue of  $\bar{\mathcal{L}}$  with associated normalized eigenvector the constant function equal to  $e_0 = \sigma(\mathbb{S}^d)^{-1/2}$ .*
2. *There exists  $\mathfrak{g} > 0$  such that for all  $\varphi \in L^2(\mathbb{S}^d)$ , the semigroup*

$$P_t\varphi(k) := \mathbb{E}_k[\varphi(m(t))]$$

*satisfies for all  $t \geq 0$*

$$\|P_t\varphi - \langle \varphi, \mathbf{1} \rangle_{L^2(\tilde{\sigma})}\|_{L^2(\mathbb{S}^d)} \leq e^{-\mathfrak{g}t} \|\varphi\|_{L^2(\mathbb{S}^d)} \quad (31)$$

*with*

$$\tilde{\sigma} = \frac{\sigma}{\sigma(\mathbb{S}^d)}$$

*the uniform measure on  $\mathbb{S}^d$ .*

3.  *$\tilde{\sigma}$  is the unique invariant measure for the Markov process  $\mathbf{M}$ .*

*Proof.* The proof of the first item is straightforward owing Lemma 4.1. Regarding the second item, we first remark that  $\tilde{\sigma}$  is an invariant measure for  $\mathbf{M}$ . This follows indeed from the fact that

$$\int_{\mathbb{S}^d} \bar{\mathcal{L}}\varphi(k) d\tilde{\sigma}(k) = 0, \quad \forall \varphi \in \mathcal{C}^\infty(\mathbb{S}^d),$$

and [11, Proposition 9.2 pp. 239]. Let now  $v(t) = \|P_t\phi\|_{L^2(\tilde{\sigma})}^2$  with  $\phi = \varphi - \langle \varphi, \mathbf{1} \rangle_{L^2(\tilde{\sigma})}$  and  $\varphi \in \mathcal{D}(\bar{\mathcal{L}})$ . Since  $\tilde{\sigma}$  is an invariant measure,  $\langle P_t\phi, \mathbf{1} \rangle_{L^2(\tilde{\sigma})} = \langle \phi, \mathbf{1} \rangle_{L^2(\tilde{\sigma})} = 0$ . Hence,

$$v'(t) = 2\langle P_t\phi, \bar{\mathcal{L}}P_t\phi \rangle_{L^2(\tilde{\sigma})} \leq -\mathfrak{g}\|P_t\phi\|_{L^2(\tilde{\sigma})}^2 = -\mathfrak{g}v(t),$$

by application of item (iii) of Lemma 4.1 to  $P_t\phi$ . Therefore,  $t \mapsto v(t)e^{\mathfrak{g}t}$  is a nonincreasing function, and then

$$v(t) = \|P_t\varphi - \langle \varphi, \mathbf{1} \rangle_{L^2(\tilde{\sigma})}\|_{L^2(\tilde{\sigma})}^2 \leq e^{-\mathfrak{g}t} \|\phi\|_{L^2(\tilde{\sigma})}^2 \leq e^{-\mathfrak{g}t} \|\varphi\|_{L^2(\tilde{\sigma})}^2.$$

The latter estimate is then extended to all  $\varphi \in L^2(\mathbb{S}^d)$  by density. For the last point, if  $\mu$  is another invariant measure, we have by definition  $\int \varphi d\mu = \int P_t\varphi d\mu$  for all  $t \geq 0$  and all  $\varphi \in \mathcal{C}^\infty(\mathbb{S}^d)$ . Because of (31), there exists  $(t_n)_n$  such that  $t_n \rightarrow +\infty$  and

$$\lim_{n \rightarrow +\infty} P_{t_n}\varphi = \int_{\mathbb{S}^d} \varphi d\tilde{\sigma} \quad \text{a.e. on } \mathbb{S}^d.$$

Dominated convergence then yields

$$\int_{\mathbb{S}^d} \varphi d\mu = \int_{\mathbb{S}^d} \varphi d\tilde{\sigma},$$

which shows that  $\mu = \tilde{\sigma}$  since  $\varphi$  is arbitrary.  $\square$

**Step 2: convergence of  $Y_x^\varepsilon$ .** We apply here standard diffusion-approximation theorems based on the martingale formulation and perturbed test functions, see e.g. [12, 20, 21, 25]. These theorems hold for instance when  $m$  is an ergodic Markov process. This is the case here since a consequence of the spectral gap estimate (31) is that  $m$  has a unique invariant measure (see Lemma 5.1 (3)), which implies that  $m$  is ergodic. Note that  $m$  is not stationary since its initial condition is not drawn according to the invariant measure. This has no consequence since (31) shows that  $\bar{\mathcal{L}}$  satisfies the Fredholm alternative, which is what is mostly needed for the construction of the perturbed test functions in the martingale techniques: as soon as  $\int_{\mathbb{S}^d} d\tilde{\sigma}(k)g(k) = 0$ , the Poisson equation

$$\bar{\mathcal{L}}u = g, \quad g \in L^2(\mathbb{S}^d)$$

admits a unique solution (up to a constant) that reads

$$u(k) = - \int_0^{+\infty} P_t g(k) dt. \quad (32)$$

Here  $(P_t)_{t \geq 0}$  is the semi-group associated with the Markov process  $\mathbf{M}$ . The integral in (32) is well defined thanks to (31).

As a result, we can apply the techniques of [12, Chapter 6] or of [14, Theorem 6.1], in order to show that  $(Y_x^\varepsilon)_\varepsilon$  converges in law in  $\mathcal{C}^0([0, \infty), \mathbb{R}^{d+1})$  as  $\varepsilon \rightarrow 0$  to a diffusion process with generator

$$\tilde{\mathcal{L}}_0 \varphi = \sum_{j,l=1}^{d+1} \partial_{x_j x_l}^2 \varphi \int_0^{+\infty} \mathbb{E}_{\tilde{\sigma}}[m_j(0)m_l(u)] du,$$

where  $\mathbb{E}_{\tilde{\sigma}}$  denotes expectation with respect to the invariant measure  $\tilde{\sigma}$ , and  $m_l(u)$  is the  $l$ -th component of the vector  $m(u)$ . The explicit form of the diffusion matrix is given in the lemma below:

**Lemma 5.2** *We have*

$$D_{jl} := \int_0^{+\infty} \mathbb{E}_{\tilde{\sigma}}[m_j(0)m_l(u)] du = \frac{1}{\mathfrak{C}} \int_{\mathbb{S}^d} d\tilde{\sigma}(k) k_j k_l,$$

with

$$\mathfrak{C} = \sigma(\mathbb{S}^{d-1}) \int_{-1}^1 F(s)(1-s^2)^{(d-2)/2}(1-s) ds \in (0, \infty).$$

*Proof.* First, let us remark that, since  $m_j(0) = k_j$ ,

$$\mathbb{E}_{\tilde{\sigma}}[m_j(0)m_l(u)] = \int_{\mathbb{S}^d} d\tilde{\sigma}(k) k_j \mathbb{E}_{\tilde{\sigma}}[m_l(u)].$$

Then, since  $m$  is a Markov process, the process  $\varphi_l(m(t)) - \varphi_l(m(0)) - \int_0^t \bar{\mathcal{L}}\varphi_l(m(u)) du$  with  $\varphi_l(p) = p_l$ , is a martingale. Hence,

$$\begin{aligned} \mathbb{E}_{\tilde{\sigma}}[m_l(t)] &= k_l + \mathbb{E}_{\tilde{\sigma}} \left[ \int_0^t \text{p.v.} \int_{\mathbb{S}^d} d\sigma(p) F(p \cdot m(u)) (p_l - m_l(u)) du \right] \\ &= k_l + \mathbb{E}_{\tilde{\sigma}} \left[ \int_0^t \lim_{\eta \rightarrow 0} \int_{-1}^{1-\eta} ds \int_{\mathbb{S}^{d-1}} d\sigma(v) F(s)(1-s^2)^{(d-2)/2} (\sqrt{1-s^2} v_l + (s-1)m_l(u)) du \right]. \end{aligned}$$

Since  $\int_{\mathbb{S}^{d-1}} d\sigma(v)v_l = 0$ , and  $F(s)(1-s^2)^{(d-2)/2}(1-s)$  is integrable thanks to (3), the limit  $\eta \rightarrow 0$  is finite and we obtain

$$\mathbb{E}_{\tilde{\sigma}}[m_l(t)] = k_l - \sigma(\mathbb{S}^{d-1}) \int_{-1}^1 ds F(s)(1-s^2)^{(d-2)/2}(1-s) \int_0^t \mathbb{E}_{\tilde{\sigma}}[m_l(u)] du.$$

Therefore, we have  $\mathbb{E}_{\tilde{\sigma}}[m_l(t)] = k_l e^{-\mathfrak{C}t}$  and then

$$D_{jl} = \frac{1}{\mathfrak{C}} \int_{\mathbb{S}^d} d\tilde{\sigma}(k) k_j k_l.$$

□

**Step 3: convergence to the diffusion equation and conclusion.** The following result shows that the limit of  $f^\varepsilon$  is characterized by the limit of  $Y_x^\varepsilon$ .

**Lemma 5.3** *We have for all  $t > 0$ , all  $f^0 \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , and all  $\varphi \in L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ :*

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^{d+1} \times \mathbb{S}^d} dx d\sigma(k) \left( f^\varepsilon(t, x, k) - \tilde{f}^\varepsilon(t, x, k) \right) \varphi(x, k) = 0,$$

where

$$\tilde{f}^\varepsilon(t, x, k) = \int_{\mathbb{S}^d} d\tilde{\sigma}(p) \mathbb{E}_k[f^0(Y_x^\varepsilon(t), p)],$$

with  $\tilde{\sigma}$  the uniform measure on  $\mathbb{S}^d$  and we recall that

$$Y_x^\varepsilon(t) = x - \varepsilon \int_0^{t/\varepsilon^2} m(u) du.$$

*Proof.* The proof follows once more from the spectral gap estimate. Following Theorem 3.1, we know that

$$f_\varepsilon(t, x, k) = T_t^\varepsilon f^0(x, k) = \mathbb{E}_k \left[ f^0 \left( x - \varepsilon \int_0^{t/\varepsilon^2} m(u) du, m(t/\varepsilon^2) \right) \right],$$

as well as  $\tilde{f}^\varepsilon$  are smooth functions (note that we defined  $T_t^\varepsilon$  on  $L^2$  here). Hence, setting

$$\phi(x, k) := f^0(x, k) - \langle f^0(x, \cdot), \mathbf{1} \rangle_{L^2(\tilde{\sigma})}, \quad \psi^\varepsilon(t, \cdot, \cdot) := T_t^\varepsilon \phi \equiv f^\varepsilon(t, \cdot, \cdot) - \tilde{f}^\varepsilon(t, \cdot, \cdot),$$

and defining  $v(t) = \|\psi^\varepsilon(t)\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}^2$ , we can differentiate  $v(t)$  and use both the transport equation (5) and item (iii) of Lemma 4.1 to arrive at

$$\begin{aligned} \|\psi^\varepsilon(t) - \langle \psi^\varepsilon(t), \mathbf{1} \rangle_{L^2(\tilde{\sigma})} \mathbf{1}\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}^2 &\leq v(t) = v(0) + \frac{2}{\varepsilon^2} \int_0^t \langle \psi^\varepsilon(s), \bar{\mathcal{L}} \psi^\varepsilon(s) \rangle_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)} ds \\ &\leq v(0) - \frac{C}{\varepsilon^2} \int_0^t \|\psi^\varepsilon(s) - \langle \psi^\varepsilon(s), \mathbf{1} \rangle_{L^2(\tilde{\sigma})} \mathbf{1}\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}^2 ds, \end{aligned}$$

which yields,

$$\|\psi^\varepsilon(t) - \langle \psi^\varepsilon(t), \mathbf{1} \rangle_{L^2(\tilde{\sigma})} \mathbf{1}\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)} \leq e^{-Ct/\varepsilon^2} \|\phi\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)}. \quad (33)$$

This shows that  $\psi^\varepsilon$  converges strongly to its average in  $k$ . We show now that this average converges weakly to zero, which implies that  $\psi^\varepsilon$  converges weakly to zero as well. Let indeed  $u^\varepsilon = \psi^\varepsilon - \langle \psi^\varepsilon, \mathbf{1} \rangle_{L^2(\tilde{\sigma})}$ , and define

$$\Psi^\varepsilon = \langle \psi^\varepsilon, \mathbf{1} \rangle_{L^2(\tilde{\sigma})}, \quad \Psi^\varepsilon(t=0) = 0,$$

which satisfies

$$\varepsilon \partial_t \Psi^\varepsilon = - \int_{\mathbb{S}^d} k \cdot \nabla_x \psi^\varepsilon d\tilde{\sigma}(k) = - \int_{\mathbb{S}^d} k \cdot \nabla_x u^\varepsilon d\tilde{\sigma}(k)$$

since  $\int_{\mathbb{S}^d} k \tilde{\sigma}(k) = 0$ . It then follows from (33), for all  $t > 0$  and for all  $\varphi \in \mathcal{C}_c^1(\mathbb{R}^{d+1})$ ,

$$\lim_{\varepsilon \rightarrow 0} \left| \int_{\mathbb{R}^{d+1}} \Psi^\varepsilon(t, x) \varphi(x) dx \right| \leq C \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} \|\nabla_x \varphi\|_{L^2(\mathbb{R}^{d+1})} \int_0^t \|u^\varepsilon(s)\|_{L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)} ds = 0. \quad (34)$$

The proof is concluded by using the density of  $\mathcal{C}_c^1(\mathbb{R}^{d+1})$  in  $L^2(\mathbb{R}^{d+1})$ , and the fact that the  $L^2$  norm in  $x$  and  $k$  of  $\psi^\varepsilon$  is uniformly bounded in  $t$  and  $\varepsilon$ .  $\square$

It thus only remains now to address the convergence of  $\tilde{f}^\varepsilon$ . When  $f^0 \in \mathcal{C}_b^0(\mathbb{R}^{d+1} \times \mathbb{S}^{d+1}) \cap L^2(\mathbb{R}^{d+1} \times \mathbb{S}^{d+1})$ , we conclude from dominated convergence and the convergence in law of  $Y_x^\varepsilon$  to  $Y_x$  that for all  $(t, x, k) \in (0, +\infty) \times \mathbb{R}^{d+1} \times \mathbb{S}^d$ ,  $\tilde{f}_\varepsilon(t, x, k)$  converges pointwise to

$$\tilde{f}(t, x, k) = \int_{\mathbb{S}^d} d\tilde{\sigma}(p) \mathbb{E}^{Y_x} [f^0(y_t, p)],$$

where we recall that  $\mathbb{E}^{Y_x}$  denotes expectation with respect to the law of  $Y_x$ . Note that this limiting law is independent of  $k$  since the diffusion coefficient  $D$  is itself independent of  $k$ , and we can write

$$\tilde{f}(t, x, k) = \tilde{f}(t, x) = \mathbb{E}^{Y_x} [\tilde{f}^0(y_t)], \quad \text{where} \quad \tilde{f}^0(x) = \int_{\mathbb{S}^d} d\tilde{\sigma}(p) f^0(x, p).$$

The latter form is a probabilistic representation of the unique solution to the diffusion equation  $\partial_t u = \tilde{\mathcal{L}}_0 u$  with initial condition  $u(0, x) = \tilde{f}^0(x)$ , and therefore  $\tilde{f}(t, x)$  is the unique solution to (6). When  $f^0$  is only an  $L^2$  function, we consider a regularized version of it denoted by  $\tilde{f}_n^0$ . As above, the associated  $\tilde{f}_n^\varepsilon$  converges pointwise in  $(t, x)$  to  $\mathbb{E}^{Y_x} [\tilde{f}_n^0(y_t)]$ . It is then not difficult to show that solutions to (6) satisfy an estimate of the form  $\|f(t)\|_{L^2} \leq \|f(0)\|_{L^2}$ , and this, together with the latter pointwise convergence and the fact that the semigroup  $T_t^\varepsilon$  is continuous in  $L^2$  show that, for all  $t > 0$ ,

$$\int_{\mathbb{R}^{d+1} \times \mathbb{S}^d} dx d\sigma f^\varepsilon(t, x, k) \varphi(x, k) \rightarrow \int_{\mathbb{R}^{d+1} \times \mathbb{S}^d} dx d\sigma \tilde{f}(t, x) \varphi(x, k), \quad \forall \varphi \in \mathcal{C}_c^0(\mathbb{R}^{d+1} \times \mathbb{S}^d).$$

Since  $\mathcal{C}_c^0(\mathbb{R}^{d+1} \times \mathbb{S}^d)$  is dense in  $L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , this implies that  $f^\varepsilon(t, x, k)$  converges, for all  $t > 0$ , weakly in  $L^2$  to  $f(t, x)$ . This concludes the proof of Theorem 3.2.

## 6 Proof of Theorem 3.3

The convergence of  $(f_\varepsilon)_\varepsilon$ , as well as the one of the measure  $\mathbb{X}_{x,k}^\varepsilon$ , is obtained via the convergence of the law of  $m^\varepsilon$  starting at  $k$ , that we denote by  $\mathbb{M}_k^\varepsilon$ . In fact, we have

$$f_\varepsilon(t, x, k) = \mathbb{E}_k \left[ f^0 \left( x - \int_0^t m^\varepsilon(u) du, m^\varepsilon(t) \right) \right] = \mathbb{E}^{\mathbb{M}_k^\varepsilon} [f^0(\Theta_x(t))],$$



where

$$\begin{aligned}\Theta_x : \mathcal{D}([0, +\infty), \mathbb{S}^d) &\longrightarrow \mathcal{D}([0, +\infty), \mathbb{R}^{d+1} \times \mathbb{S}^d) \\ \omega &\mapsto t \mapsto \left(x - \int_0^t \omega(u) du, \omega(t)\right)\end{aligned}$$

is a continuous application for the Skorohod topology. Note that we have  $\mathbb{X}_{x,k}^\varepsilon = \mathbb{M}_k^\varepsilon \circ \Theta_x^{-1}$ .

**Lemma 6.1**  $(\mathbb{M}_k^\varepsilon)_\varepsilon$  is tight in  $\mathcal{D}([0, +\infty), \mathbb{S}^d)$ .

*Proof.* We will prove this lemma by using Aldous's tightness criterion [8, Theorem 16.10 pp. 178]. The first requirement is

$$\lim_{M \rightarrow +\infty} \sup_{\varepsilon \in (0,1)} \mathbb{M}_k^\varepsilon \left( \sup_{0 \leq t \leq T} \|y_t\|_{\mathbb{R}^{d+1}} \geq M \right) = 0 \quad \forall T > 0,$$

which is direct since  $\mathbb{M}_k^\varepsilon$  is supported on  $\mathbb{S}^d$ . Now, let  $T > 0$  and  $\tau$  be a discrete stopping time relatively to the canonical filtration, and bounded by  $T$ . Considering also two real numbers  $\mu > 0$  and  $\nu > 0$ , we have

$$\mathbb{M}_k^\varepsilon(\|y_{\tau+\mu} - y_\tau\|_{\mathbb{R}^{d+1}} > \nu) \leq \sum_{j=1}^{d+1} \mathbb{M}_k^\varepsilon(|y_{\tau+\mu}^j - y_\tau^j| > \nu/(d+1)),$$

where  $y^j$  stands for the  $j$ -th component of  $y$ . In order to study the increment, let us introduce

$$\begin{aligned}M_j(t) &= y_t^j - \left( \int_0^t du \bar{\mathcal{L}}_\varepsilon(p - y_u) \right)_j \\ &= y_t^j - \int_0^t du \text{ p.v. } \int_{\mathbb{S}^d} d\sigma(p) F^\varepsilon(p \cdot y_u)(p_j - y_u^j)\end{aligned}$$

which is a martingale under  $\mathbb{M}_k^\varepsilon$  (see [11, Proposition 1.7 pp. 162]) with quadratic variation

$$\langle M_j \rangle (t) = \int_0^t (\bar{\mathcal{L}}_\varepsilon \varphi^2(y_u^j) - 2\varphi(y_u^j) \bar{\mathcal{L}}_\varepsilon \varphi(y_u^j)) du, \quad (35)$$

where  $\varphi(x) = x$ . Hence,

$$\begin{aligned}&\mathbb{M}_k^\varepsilon(|y_{\tau+\mu}^j - y_\tau^j| > \nu/(d+1)) \\ &\leq \mathbb{M}_k^\varepsilon(|M_j(\tau + \mu) - M_j(\tau)| > \nu/(2(d+1))) \\ &+ \mathbb{M}_k^\varepsilon\left(\left| \int_\tau^{\tau+\mu} du \text{ p.v. } \int_{\mathbb{S}^d} d\sigma(p) F^\varepsilon(p \cdot y_u)(p_j - y_u^j) \right| > \nu/(2(d+1))\right) \\ &\leq I + II.\end{aligned}$$

Regarding the second term, proceeding as in the proof of Lemma 5.2, we find

$$\left| \int_\tau^{\tau+\mu} du \text{ p.v. } \int_{\mathbb{S}^d} d\sigma(p) F^\varepsilon(p \cdot y_u)(p_j - y_u^j) \right| \leq \mathfrak{C} \int_\tau^{\tau+\mu} du |y_u^j| \leq \mu \mathfrak{C}, \quad (36)$$

for all  $\mu > 0$ , and therefore  $II = 0$ . Now, regarding the term  $I$ , we have using the martingale property of  $M_j$

$$\begin{aligned} \mathbb{M}_k^\varepsilon(|M_j(\tau + \mu) - M_j(\tau)|^2 > \nu^2/(4(d+1)^2)) \\ \leq \frac{4(d+1)^2}{\nu^2} \mathbb{E}^{\mathbb{M}_k^\varepsilon}[|M_j(\tau + \mu) - M_j(\tau)|^2] \\ \leq \frac{4(d+1)^2}{\nu^2} \mathbb{E}^{\mathbb{M}_k^\varepsilon}[\langle M_j \rangle(\tau + \mu) - \langle M_j \rangle(\tau)] \\ \leq C_{2,\nu}\mu, \end{aligned}$$

where we used (35) and similar arguments as in the derivation of (36). As a result, we obtain

$$\lim_{\mu \rightarrow 0} \sup_{\varepsilon} \sup_{\tau} \mathbb{M}_k^\varepsilon(\|y_{\tau+\mu} - y_\tau\|_{\mathbb{R}^{d+1}} > \nu) = 0,$$

which concludes the proof of Lemma 6.1.  $\square$

We will use the following lemma that shows the convergence of the generators.

**Lemma 6.2** *We have for all  $\varphi \in \mathcal{C}^\infty(\mathbb{S}^d)$ ,*

$$\lim_{\varepsilon \rightarrow 0} \|\mathcal{L}_\varepsilon \varphi - \mathcal{L}_\beta \varphi\|_{L^\infty(\mathbb{S}^d)} = 0.$$

*Proof.* Let  $\delta F^\varepsilon(s) := \varepsilon^{\beta+d/2} K(\varepsilon(1-s)) - a_1|1-s|^{-\beta-d/2}$ . According to (7), for all  $s \in (-1, 1)$  and for all  $\mu > 0$ , there exists  $\varepsilon_0 > 0$  such that,  $\forall \varepsilon \in (0, \varepsilon_0)$ ,

$$|\delta F^\varepsilon(s)| \leq \mu |1-s|^{-\beta-d/2}. \quad (37)$$

We then write

$$\delta \mathcal{L}_\varepsilon \varphi := \mathcal{L}_\varepsilon \varphi - \mathcal{L}_\beta \varphi = \text{p.v.} \int_{\mathbb{S}^d} d\sigma(p) \delta F^\varepsilon(p \cdot k) (\varphi(p) - \varphi(k)).$$

We may assume without loss of generality that  $k = e_{d+1} = (0, \dots, 0, 1)$ , and write

$$p = (\sqrt{1-s^2}u, s),$$

with  $s \in [-1, 1]$ , and  $u \in \mathbb{S}^{d-1}$ . Then,

$$\delta \mathcal{L}_\varepsilon \varphi(k) = \lim_{\eta \rightarrow 0} \int_{\mathbb{S}^{d-1}} \int_{-1}^{1-\eta} \delta F_\varepsilon(s) (\varphi(\sqrt{1-s^2}u + sk) - \varphi(k)) (1-s^2)^{\frac{d-2}{2}} d\sigma(u) ds.$$

Recasting  $\varphi$  as

$$\varphi(k) = \phi(0, \dots, 0, 1), \quad \varphi(\sqrt{1-s^2}u + sk) = \phi(\sqrt{1-s^2}u_1, \dots, \sqrt{1-s^2}u_d, s),$$

we have

$$\begin{aligned} \varphi(\sqrt{1-s^2}u + sk) - \varphi(k) &= (s-1)\partial_{x_{d+1}}\phi(0, \dots, 0, 1) + \sqrt{1-s^2}u \cdot \nabla_d \phi(0, \dots, 0, 1) \\ &\quad + \mathcal{O}(|s-1|), \end{aligned}$$

where  $\partial_{x_{d+1}}$  denotes partial derivation with respect to the  $(d+1)$ -th variable and  $\nabla_d$  the gradient with respect to the  $d$  first variables. Since  $\int_{\mathbb{S}^{d-1}} u d\sigma(u) = 0$ , it follows from the equation above and (37) that

$$|\delta \mathcal{L}_\varepsilon f(k)| \leq C\mu \int_{-1}^1 \frac{(1-s)(1-s^2)^{\frac{d-2}{2}}}{(1-s)^{\beta+\frac{d}{2}}} ds \leq C\mu \int_{-1}^1 \frac{ds}{(1-s)^\beta} \leq C\mu, \quad (38)$$

since  $\beta \in (0, 1)$ . This concludes the proof of the lemma.  $\square$

We have all needed now to conclude the proof. We deduce from Lemma 6.1 that, up to the extraction of a subsequence,  $\mathbb{M}_k^\varepsilon$  converges weakly to a measure  $\mathbb{M}_k$ . As a consequence of Lemma 6.2,  $\mathbb{M}_k$  is a solution to the martingale problem associated to  $\bar{\mathcal{L}}_\beta$  and starting at  $k$ . Since the equation  $\partial_t u = \bar{\mathcal{L}}_\beta u$  admits a unique solution for a given initial condition, it turns out that this martingale problem is well-posed (c.f. [11, Theorem 4.2 pp. 184]), so that the entire sequence converges since the limiting measure is unique.

When  $f_0 \in \mathcal{C}_b^0(\mathbb{R}^{d+1} \times \mathbb{S}^d) \cap L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , then, pointwise in  $(t, x, k)$ ,

$$f_\varepsilon(t, x, k) = \mathbb{E}^{\mathbb{M}_k^\varepsilon \circ \Theta_x^{-1}} [f^0(y_t)] \rightarrow \mathbb{E}^{\mathbb{X}_{x,k}} [f^0(y_t)],$$

where  $\mathbb{X}_{x,k} = \mathbb{M}_k \circ \Theta_x^{-1}$  is the law of a diffusion process with generator  $-k \cdot \nabla_x + \bar{\mathcal{L}}_\beta$  and starting at  $(x, k)$ . This, together with Lemma 6.2 and dominated convergence, allows us to pass to the limit in the weak formulation of (8), and to deduce that the limit above is a weak solution to (9). According to Theorem 3.1, the solution to (9) is unique and actually  $\mathcal{C}^\infty$ .

When  $f_0$  is only in  $L^2(\mathbb{R}^{d+1} \times \mathbb{S}^d)$ , we use a regularization procedure similar to the one at the end of the proof of Theorem 3.2. This concludes the proof of Theorem 3.3.

## References

- [1] R. ALEXANDRE, *Fractional order kinetic equations and hypoellipticity*, Anal. Appl. (Singap.), 10 (2012), pp. 237–247.
- [2] R. ALEXANDRE, Y. MORIMOTO, S. UKAI, C.-J. XU, AND T. YANG, *Uncertainty principle and kinetic equations*, J. Funct. Anal., 255 (2008), pp. 2013–2066.
- [3] ———, *Global existence and full regularity of the Boltzmann equation without angular cutoff*, Comm. Math. Phys., 304 (2011), pp. 513–581.
- [4] R. ALEXANDRE AND C. VILLANI, *On the Boltzmann equation for long-range interactions*, Comm. Pure Appl. Math., 55 (2002), pp. 30–70.
- [5] R. ALONSO AND W. SUN, *The radiative transfer equation in the forward-peaked regime*, Comm. Math. Phys., 338(3) (2015), pp. 1233–1286.
- [6] K. ATKINSON AND W. HAN, *Spherical harmonics and approximations on the unit sphere: an introduction*, vol. 2044 of Lecture Notes in Mathematics, Springer, Heidelberg, 2012.
- [7] G. BAL, T. KOMOROWSKI, AND L. RYZHIK, *Kinetic limits for waves in a random medium*, Kinet. Relat. Models, 3 (2010), pp. 529–644.
- [8] P. BILLINGSLEY, *Convergence of Probability Measures*, John Wiley and Sons, New York, 1999.
- [9] F. BOUCHUT, *Hypoelliptic regularity in kinetic equations*, J. Math. Pures Appl. (9), 81 (2002), pp. 1135–1159.

- [10] L. ERDÖS AND H. T. YAU, *Linear Boltzmann equation as scaling limit of quantum Lorentz gas*, Advances in Differential Equations and Mathematical Physics. Contemporary Mathematics, 217 (1998), pp. 137–155.
- [11] S. N. ETHIER AND T. G. KURTZ, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986. Characterization and convergence.
- [12] J.-P. FOUQUE, J. GARNIER, G. PAPANICOLAOU, AND K. SÖLNA, *Wave propagation and time reversal in randomly layered media*, vol. 56 of Stochastic Modelling and Applied Probability, Springer, New York, 2007.
- [13] M. FUKUSHIMA, *Dirichlet forms and Markov processes*, vol. 23 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1980.
- [14] J. GARNIER, *A multi-scaled diffusion-approximation theorem. applications to wave propagation in random media*, ESAIM Probab. Statist., 1 (1997), pp. 183–206.
- [15] P. GÉRARD, P. A. MARKOWICH, N. J. MAUSER, AND F. POUPAUD, *Homogenization limits and Wigner transforms*, Comm. Pure Appl. Math., 50 (1997), pp. 323–380.
- [16] C. GOMEZ, *Radiative transport limit for the random Schrödinger equation with long-range correlations*, J. Math. Pures Appl. (9), 98 (2012), pp. 295–327.
- [17] C. GOMEZ, O. PINAUD, AND L. RYZHIK, *Hypoelliptic estimates in radiative transfer*, to appear in Commun. Part. Diff. Eq.
- [18] P. T. GRESSMAN AND R. M. STRAIN, *Global classical solutions of the Boltzmann equation with long-range interactions*, Proc. Natl. Acad. Sci. USA, 107 (2010), pp. 5744–5749.
- [19] P. T. GRESSMAN AND R. M. STRAIN, *Global classical solutions of the Boltzmann equation without angular cut-off*, J. Amer. Math. Soc., 24 (2011), pp. 771–847.
- [20] H. KUNITA, *Stochastic flows and stochastic differential equations*, vol. 24 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1990.
- [21] H. J. KUSHNER, *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory*, MIT Press Series in Signal Processing, Optimization, and Control, 6, MIT Press, Cambridge, MA, 1984.
- [22] N. LERNER, Y. MORIMOTO, AND K. PRAVDA-STAROV, *Hypoelliptic estimates for a linear model of the Boltzmann equation without angular cutoff*, Comm. Partial Differential Equations, 37 (2012), pp. 234–284.
- [23] P.-L. LIONS AND T. PAUL, *Sur les mesures de Wigner*, Rev. Mat. Iberoamericana, 9 (1993), pp. 553–618.
- [24] J. LUKKARINEN AND H. SPOHN, *Kinetic limit for wave propagation in a random medium*, Arch. Ration. Mech. Anal., 183 (2007), pp. 93–162.
- [25] 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), Paper No. 6, Duke Univ., Durham, N.C., 1977, pp. ii+120 pp. Duke Univ. Math. Ser., Vol. III.
- [26] M. REED AND B. SIMON, *Methods of modern mathematical physics. II. Fourier Analysis, Self-Adjointness*, Academic Press, Inc., New York, second ed., 1980.
- [27] S. SAMKO, *On inversion of fractional spherical potentials by spherical hypersingular operators*, in Singular integral operators, factorization and applications, vol. 142 of Oper. Theory Adv. Appl., Birkhäuser, Basel, 2003, pp. 357–368.
- [28] H. SPOHN, *Derivation of the transport equation for electrons moving through random impurities*, J. Stat. Phys., 17 (1977), pp. 385–412.
- [29] F. TRÈVES, *Topological vector spaces, distributions and kernels*, Academic Press, New York, 1967.

- [30] E. WIGNER, *On the quantum correction for thermodynamic equilibrium*, Phys. Rev. 40, (1932), pp. 742–749.