THE BROWNIAN MOTION AS THE LIMIT OF A DETERMINISTIC SYSTEM OF HARD-SPHERES

THIERRY BODINEAU, ISABELLE GALLAGHER AND LAURE SAINT-RAYMOND

Abstract. We provide a rigorous derivation of the brownian motion as the hydrodynamic limit of systems of hard-spheres as the number of particles N goes to infinity and their diameter ε simultaneously goes to 0, in the fast relaxation limit $N\varepsilon^{d-1} \to \infty$ (with a suitable scaling of the observation time and length).
As suggested by Hilbert in his sixth problem, we use the linear Boltzmann equation as an intermediate level of description for one tagged particle in a gas close to global equilibrium. Our proof relies on the fundamental ideas of Lanford. The main novelty here is the detailed study of the branching process, leading to explicit estimates on pathological collision trees.

1. Introduction

The sixth problem raised by Hilbert in 1900 on the occasion of the International Congress of Mathematicians addresses the question of the axiomatization of mechanics, and more precisely of describing the transition between atomistic and continuous models for gas dynamics by some mathematical convergence results.

Even though it is quite restrictive (since only perfect gases can be considered by this process), Hilbert further suggested to use Boltzmann's kinetic equation as an intermediate step to understand the appearance of irreversibility and dissipative mechanisms [18]:
"*Quant aux principes de la Mécanique, nous possédons déjà au point de vue physique des recherches d'une haute portée; je citerai, par exemple, les écrits de MM. Mach, Hertz, Boltzmann et Volkmann. Il serait aussi très désirable qu'un examen approfondi des principes de la Mécanique fût alors tenté par les mathématiciens. Ainsi le Livre de M. Boltzmann sur les Principes de la Mécanique nous incite à établir et à discuter au point de vue mathématique d'une manière complète et rigoureuse les méthodes basées sur l'idée de passage à la limite, et qui de la conception atomique nous conduisent aux lois du mouvement des continua. Inversement on pourrait, au moyen de méthodes basées sur l'idée de passage à la limite, chercher à déduire les lois du mouvement des corps rigides d'un système d'axiomes reposant sur la notion d'états d'une matière remplissant tout l'espace d'une manière continue, variant d'une manière continue et que l'on devra définir paramétriquement.*

Quoi qu'il en soit, c'est la question de l'équivalence des divers systèmes d'axiomes qui présentera toujours l'intérêt le plus grand quant aux principes."

A huge literature has been devoted to these asymptotic problems, but up to now they remain still largely open. Important breakthroughs [13, 2] have allowed for a complete study of some hydrodynamic limits of the Boltzmann equation, especially in incompressible viscous regimes leading to the Navier-Stokes equations (see [17] for instance). Note that other regimes such as the compressible Euler limit (which is the most immediate from the formal point of view) are still far from being understood.
But, at this stage, the main obstacle seems actually to come from the other step, namely from the derivation of the Boltzmann equation from a system of interacting particles: the

Date: May 16, 2013.

best result concerning this low density limit which is due to Lanford in the case of hard-spheres [22] (see also [9, 33, 14] for a complete proof) is indeed valid only for short times, i.e. breaks down before any relaxation can be observed.

Theorem 1.1. *Consider a system of N particles interacting*

- *either as hard-spheres of diameter ε*
- *or via a repulsive potential Φ_ε, with compact support $B(0,\varepsilon)$, radial and singular at 0, and such that the scattering of particles can be parametrized by their deviation angle.*

Let $f_0 : \mathbf{R}^{2d} \mapsto \mathbf{R}^+$ be a continuous density of probability such that

$$\big\| f_0 \exp(\frac{\beta}{2}|v|^2)\big\|_{L^\infty(\mathbf{R}^d_x\times\mathbf{R}^d_v)} < +\infty$$

for some $\beta > 0$.
Assume that the N particles are initially distributed according to f_0 and "independent". Then, there exists some $\tau > 0$ such that, in the Boltzmann-Grad limit $N \to \infty$, $N\varepsilon^{d-1} = 1$, the distribution function of the particles converges in the sense of observables (i.e. averages in v) to the solution of the Boltzmann equation

$$\begin{aligned}
&\partial_t f + v\cdot\nabla_x f = Q(f,f),\\
&Q(f,f)(v) := \iint_{\mathbf{S}^{d-1}\times\mathbf{R}^d} [f(v')f(v_1') - f(v)f(v_1)]\, b(v-v_1,\nu)\, dv_1 d\nu \qquad (1.1)\\
&v' = v + \nu\cdot(v_1-v)\,\nu\,, \quad v_1' = v_1 - \nu\cdot(v_1-v)\,\nu\,,
\end{aligned}$$

with a locally bounded cross-section b depending on Φ_ε implicitly, and with initial data f_0. In the case of a hard-sphere interaction, the cross section is given by

$$b(v-v_1,\nu) = \big((v-v_1)\cdot\nu\big)_+ \, .$$

Here, by "independent", we mean typically that the initial N-particle distribution is obtained by factorization and conditioning on energy surfaces (see [14] and references therein). In the case of hard-spheres for instance, one would have

$$f_{N|t=0} = \frac{1}{\mathcal{Z}_N} f_0^{\otimes N} \mathbf{1}_{\mathcal{D}^N_\varepsilon}\,,$$

with

$$\mathcal{D}^N_\varepsilon := \big\{ (x_1,v_1,\dots,x_N,v_N) \in \mathbf{R}^{2dN} \,/\, \forall i\neq j\,, \quad |x_i - x_j| > \varepsilon \big\}$$

and

$$f_0^{\otimes N}(x_1,v_1,\dots,x_N,v_N) := \prod_{i=1}^N f_0(x_i,v_i)\,,$$

while $\mathcal{Z}_N$ normalizes the integral of $f_{N|t=0}$ to 1.

The main reason why the convergence is not known to hold for longer time intervals is the fact that we are not able to rule out the possibility of spatial concentrations of the density leading to some pathological collision process.

What we intend to do here is to overcome this difficulty by considering a good notion of fluctuation around global equilibrium for the system of interacting particles, and to get a complete derivation of the hydrodynamic limit from the hard-sphere system in this linear regime. Of course, in this framework one cannot hope to retrieve a model for the full (non-linear) gas dynamics, but – as far as we know – this is the very first **result describing the Brownian motion as the limit of a deterministic system of interacting particles**.

The main difficulty here is to justify the approximation by the linear Boltzmann equation

$$
\begin{aligned}
&\partial_t f + v\cdot\nabla_x f = -Lf,\\
&Lf(v) := \iint [f(v)M_\beta(v_1) - f(v')M_\beta(v_1')]\, b(v-v_1,\nu)\, dv_1 d\nu,\\
&M_\beta(v) := \left(\frac{\beta}{2\pi}\right)^{\frac d2} \exp\left(-\frac{\beta}{2}|v|^2\right), \quad \beta>0\,,
\end{aligned}
\tag{1.2}
$$

for times diverging as $N\to\infty$. Indeed, in diffusive regime, the convergence of the linear Boltzmann process towards the Brownian is by now a classical result [21].

2. Strategy and main results

The good notion of fluctuation is obtained by considering the motion of a tagged particle (or possibly a finite set of tagged particles) in a gas of N particles at equilibrium (or close to equilibrium), in the limit $N\to\infty$.

2.1. **The Lorentz gas.**

If the background particles are infinitely heavier than the tagged particle then the dynamics can be approximated by a Lorentz gas, i.e. by the motion of the tagged particle in a frozen background. The linear Boltzmann equation has been derived (globally in time) from the dynamics of a tagged particle in a low density Lorentz gas, meaning that

- the obstacles are distributed randomly according to some Poisson law;
- the obstacles have no dynamics, in particular they do not feel the effect of collisions with the tagged particle.

This problem, suggested by Lorentz [25] at the beginning of the twentieth century to study the motion of electrons in metals, is the core of a number of works, and the corresponding literature includes a large variety of contributions. We do not intend to be exhaustive here and refer the reader to the book by Spohn [32] (Chapter 8) for a survey on this topic. We state one basic result due to Gallavotti [15] in the low density limit and then indicate some of the many important research directions.

Theorem 2.1. *Consider a Poisson distribution of fixed scatterers in $\mathbf{R}^d$, of parameter ε^{1-d}. Let $T^t_{\varepsilon,c}$ be the flow generated by this distribution of obstacles c (with hard-sphere interactions with the scatterers). For a given continuous initial datum $f^0\in L^1\cap L^\infty(\mathbf{R}^{2d})$, we define*

$$
f_N(t,x,v) := \mathbb{E}[f^0(T^{-t}_{\varepsilon,c}(x,v))]\,.
$$

Then, for any time $T>0$, f_N converges to the solution f of the linear Boltzmann equation (1.2), with hard-sphere cross-section, in $L^\infty([0,T],L^1(\mathbf{R}^{2d}))$.

A refinement of this result can be found for instance in [30] in terms of convergence of path measures (and not only of the mean density), as well as in [6] where the convergence is proven for typical scatterer configurations (and not only in average).

These convergence statements lead naturally to various questions concerning

- the assumptions on the microscopic potential of interaction,
- the role of randomness for the distribution of scatterers,
- the long time behavior of the system, in particular the relaxation towards thermodynamic equilibrium and hydrodynamic limits.

The first point has been addressed by Desvillettes, Pulvirenti and Ricci [11, 12]. Their goal was to derive “singular ” kinetic equations such as the linear Boltzmann equation without angular cut-off or the Fokker-Planck equation, from a system of particles with long-range interactions. They have obtained partial results in this direction, insofar as they can consider only asymptotically long-range interactions. Due to the fact that the range of the potential is infinite in the limit, the test particle interacts typically with infinitely many obstacles, so the set of bad configurations of the scatterers (such as the set of configurations yielding recollisions), preventing the Markov property of the limit must be estimated explicitly. Even though the long-range tails add a very small contribution to the total force for each typical scatterer distribution, the non grazing collisions generate an exponential instability making the two trajectories (with and without cut-off) very different. The complete derivation of the linear Boltzmann equation for long-range interactions is therefore still open, although the methods developed in this paper could be a first angle of attack, at least for sufficiently decaying potentials [4].

The second important research direction, initiated by Golse, is related to the distribution of scatterers. It is often more appropriate from a physical point of view to consider more general distributions of obstacles than the Poisson distribution. In particular, in the original problem of Lorentz, the atoms of metal are distributed on a periodic network. In this case, the Boltzmann-Grad limit of the Lorentz gas with a periodic distribution of scatterers cannot be described by a linear Boltzmann equation [16, 8, 26, 27].

2.2. An alternative strategy based on the maximum principle.

We adopt here a slightly different point of view, and consider a deterministic system of N hard-spheres, meaning that the tagged particle is identical to the particles of the background, interacting according to the same collision laws.

On the one hand, the problem seems more difficult insofar as the background has its own dynamics, which is coupled with the tagged particle. But, on the other hand, pathological situations as described in [16, 7, 8] are not stable: because of the dynamics of the scatterers, we expect the situation to be better since some ergodicity could be retrieved from the additional degrees of freedom. In particular, there are invariant measures for the whole system, i.e. the system consisting of both the background and the tagged particle.

Here we shall take advantage of the latter property to establish global uniform a priori bounds for the distribution of particles, and more generally for all marginals of the N-particle distribution (see Proposition 3.2). This will be the key to control the collision process, and to prove (like in Kac's model [20] for instance) that dynamics for which a very large number of collisions occur over a short time interval, are of vanishing probability.

Note that a similar strategy, based on the existence of the invariant measure, was already used by Lebowitz and Spohn [24] to derive the linear Boltzmann equation for long times.

Let us now give the precise framework of our study. As explained above, the idea is to improve the result by Lanford by considering fluctuations around some global equilibrium. Locally the N-particle distribution f_N should therefore look like a conditioned tensorized Maxwellian.

In order for the marginals of f_N to be well-defined, we shall consider the dynamics on a bounded domain $\Omega \subset \mathbf{R}^d$, the size of which may possibly tend to infinity in the limit. We shall further restrict our attention to simple geometries in order to avoid pathologies related to boundary effects, and complicated free dynamics. The simplest example is the case of a periodic domain. Let $\mathbf{T}^d_\lambda$ be the d-dimensional torus, where

$$\mathbf{T}_\lambda := \mathbf{R}/(\lambda \mathbf{Z})$$

for some $\lambda \geq 1$. With these notations, the average number of particles per unit ~~of~~ volume is

$$N' = N\lambda^{-d}$$

so that the Boltzmann-Grad scaling is given by $N\varepsilon^{d-1}\lambda^{-d} \equiv 1$, i.e.

$$N = \lfloor \frac{\lambda^d}{\varepsilon^{d-1}} \rfloor ,$$

where $\lfloor \cdot \rfloor$ stands for the integer part. Note that λ can also diverge in the limit $N \to \infty$.

In all the sequel, we shall focus on the case of the hard-sphere dynamics to avoid technicalities due to artificial boundaries and cluster estimates. We however point out that our main results still hold for the low density limit of system of interacting particles with a compactly supported interaction potential (typically satisfying the same assumptions as in [14] or in [29]). The microscopic model is therefore given by the following system of ODEs: for $1 \leq i \neq j \leq N$,

$$\begin{aligned} &\frac{dx_i}{dt} = v_i, \quad \frac{dv_i}{dt} = 0 \quad \text{as long as } |x_i(t) - x_j(t)| > \varepsilon , \\ &\left.\begin{aligned} v_i(t^+) &= v_i(t^-) - \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j)(t^-) \\ v_j(t^+) &= v_j(t^-) + \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j)(t^-) \end{aligned}\right\} \quad \text{if } |x_i(t) - x_j(t)| = \varepsilon . \end{aligned} \tag{2.1}$$

In the following we denote, for $1 \leq i \leq N$, $z_i := (x_i, v_i)$ and $Z_N := (z_1, \dots, z_N)$. With a slight abuse we say that Z_N belongs to $\mathbf{T}^{dN}_\lambda \times \mathbf{R}^{dN}$ if $X_N := (x_1, \dots, x_N)$ belongs to $\mathbf{T}^{dN}_\lambda$ and $V_N := (v_1, \dots, v_N)$ to $\mathbf{R}^{dN}$. Recalling that

$$\mathcal{D}^N_\varepsilon := \big\{ Z_N \in \mathbf{T}^{dN}_\lambda \times \mathbf{R}^{dN} \, / \, \forall i \neq j \, , \quad |x_i - x_j| > \varepsilon \big\}$$

we also define for $1 \leq i \neq j \leq N$

$$\partial\mathcal{D}^N_\varepsilon(i,j) := \Big\{ Z_N \in \mathbf{T}^{dN}_\lambda \times \mathbf{R}^{dN} \, / \, |x_i - x_j| = \varepsilon \text{ and } \forall (k,\ell) \in [1,N] \setminus \{i,j\} \, , |x_k - x_\ell| > \varepsilon \Big\} \, .$$

Given Z_N on $\partial\mathcal{D}^N_\varepsilon(i,j)$, if

$$(v_i - v_j)\cdot(x_i - x_j) < 0$$

then we define $Z^{in}_N(i,j)$ and $Z^{out}_N(i,j)$ as follows:

$$Z^{in}_N(i,j) := Z_N \quad \text{and} \quad X^{out}_N(i,j) := X_N \tag{2.2}$$

while

$$\begin{aligned} v^{out}_k(i,j) &:= v_k \quad \forall k \in [1,N] \setminus \{i,j\} \, , \\ v^{out}_i(i,j) &:= v^{in}_i - \frac{1}{\varepsilon^2}(v^{in}_i - v^{in}_j)\cdot(x_i - x_j)(x_i - x_j) \\ v^{out}_j(i,j) &:= v^{in}_j + \frac{1}{\varepsilon^2}(v^{in}_i - v^{in}_j)\cdot(x_i - x_j)(x_i - x_j) \, . \end{aligned} \tag{2.3}$$

Defining the Hamiltonian

$$H_N(Z_N) := \frac{1}{2}\sum_{i=1}^N |v_i|^2 \, ,$$

we consider the Liouville equation in the $2Nd$-dimensional phase space $\mathcal{D}^N_\varepsilon$

$$\partial_t f_N + \{H_N, f_N\} = 0 \tag{2.4}$$

with specular reflection on the boundary, meaning that if Z_N belongs to $\partial\mathcal{D}^N_\varepsilon(i,j)$ with

$$(v_i - v_j)\cdot(x_i - x_j) < 0$$

then

$$f_N(t, Z^{out}_N(i,j)) = f_N(t, Z^{in}_N(i,j))$$

with the notation introduced in (2.2)-(2.3). We recall, as shown in [1] for instance, that the set of initial configurations leading to an ill-defined problem (clustering of collision times, or collisions involving more than two particles) is of measure zero in $\mathbf{T}_\lambda^{dN} \times \mathbf{R}^{dN}$.

An obvious remark is that stationary solutions of (2.4) provide asymptotically **stationary solutions to the Boltzmann hierarchy**, such that the convergence is global in time. Any function of the energy $f_N \equiv F(H_N)$ is a stationary solution of the Liouville equation (2.4). For instance Gibbs measures are obtained by taking $F_\beta(x) = \exp(-\beta x)$: we define, for $\beta > 0$ given,

$$M_{N,\beta} := \frac{1}{\bar{\mathcal{Z}}_N} \left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N)\, \mathbf{1}_{\mathcal{D}_\varepsilon^N} \tag{2.5}$$

where the partition function $\bar{\mathcal{Z}}_N$ is nothing else than the normalization factor:

$$\begin{aligned}\bar{\mathcal{Z}}_N &:= \int_{\mathbf{T}_\lambda^{dN} \times \mathbf{R}^{dN}} \left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(Z_N)) \mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N)\, dZ_N \\ &= \int_{\mathbf{T}_\lambda^{dN}} \prod_{1 \leq i \neq j \leq N} \mathbf{1}_{|x_i - x_j| > \varepsilon}\, dX_N\,.\end{aligned} \tag{2.6}$$

Straightforward computations enable to prove asymptotically the independence.

Proposition 2.1. *Given $\beta > 0$, there is a constant $C > 0$ such that for any fixed $s \in \mathbf{N}^*$, the marginal of order s*

$$M_{N,\beta}^{(s)}(Z_s) := \int M_{N,\beta}(Z_N)\, dz_{s+1} \dots dz_N \tag{2.7}$$

satisfies, as $N \to \infty$ in the scaling $N\varepsilon^{d-1}\lambda^{-d} \equiv 1$,

$$\Big| \Big(\lambda^{ds} M_{N,\beta}^{(s)} - M_\beta^{\otimes s}\Big) \mathbf{1}_{\mathcal{D}_\varepsilon^s} \Big| \leq Cs\, \varepsilon M_\beta^{\otimes s} \tag{2.8}$$

where

$$M_\beta^{\otimes s}(Z_s) := \prod_{i=1}^s M_\beta(v_i) \quad \text{and} \quad M_\beta(v) := \left(\frac{\beta}{2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{\beta}{2}|v|^2\right)\,.$$

The proof of Proposition 2.1, by now classical, is recalled in Appendix A for the sake of completeness.

Remark 2.2. *Notice the normalization factor λ^{ds} in (2.8), which is necessary for the initial data to be asymptotic to the tensorized Maxwellian $M_\beta^{\otimes s}$. We are indeed interested in describing the local density of the gas, which is the good parameter to understand the collision process.*

In order to obtain the convergence for long times, a natural idea is to "weakly" perturb these equilibrium states $\lambda^{-dN} M_{N,\beta}$, by modifying the distribution of a finite number s of particles. In other words, we shall describe the dynamics of s tagged particles in a background at equilibrium. Following this strategy, which is classical in probability theory,

- we lose asymptotically the nonlinear coupling: we thus expect to get a linear equation for the distribution of tagged particles;
- We also lose the feedback of tagged particles on the background: since this background is constituted of $N \gg 1$ indistinguishable particles, the momentum and energy exchange with tagged particles has a very dilute effect on each one of these indistinguishable particles and does not modify the average distribution. As a consequence, the limiting equation for the distribution of tagged particles should be non conservative.

What we shall actually prove is that the limiting dynamics is governed by the linear Boltzmann equation (1.2), with hard-sphere cross-section.

2.3. **Main results.** For the sake of simplicity, we choose only one tagged particle (numbered 1) and consider initial data of the form

$$f_N^0(Z_N) := \frac{1}{\mathcal{Z}_N}\left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(Z_N))\varphi_N^0(z_1)\mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N)\,, \tag{2.9}$$

with

$$\mathcal{Z}_N := \int_{\mathbf{T}_\lambda^{dN}\times\mathbf{R}^{dN}} \left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(Z_N))\varphi_N^0(z_1)\mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N)\,dZ_N$$
$$= \left(\frac{\beta}{2\pi}\right)^{\frac{d}{2}} \int_{\mathbf{T}_\lambda^{dN}\times\mathbf{R}^{d}} \exp(-\frac{\beta}{2}|v_1|^2)\varphi_N^0(z_1) \prod_{1\leq i\neq j\leq N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\,dX_N dv_1\,.$$

We assume that there exists μ_N and Λ_N (the dependence on N of which will be prescribed later) such that uniformly in $\mathbf{T}_\lambda^d\times\mathbf{R}^d$,

$$\mu_N^{-1}\frac{1}{\xi^d}1_{|x|_\infty\leq\xi} \leq \varphi_N^0(z_1)\leq \mu_N \quad\text{and}\quad |\nabla_{z_1}\varphi_N^0(z_1)|\leq\Lambda_N\,, \tag{2.10}$$

for some constant $\xi>0$.

Remark 2.3. *The pointwise bound from below in* (2.10) *will be used only to get an estimate on $\mathcal{Z}_N^{-1}$ (see the proof of Proposition* 3.2*). The precise value of the parameter ξ in* (2.10) *is not relevant, indeed paving $\mathbf{T}_\lambda^d$ by boxes of side length ξ and using the shift invariance of the measure, we obtain easily the lower bound $\mathcal{Z}_N\geq\mu_N^{-1}\bar{\mathcal{Z}}_N$ uniformly in ξ.*

In all the sequel, given two positive parameters η_1 and η_2, we shall say that

$$\eta_1\ll\eta_2 \ \text{ if } \ \eta_1\leq C\eta_2$$

for some large constant C which does not depend on any parameter (other than β and d).

Using the global uniform a priori bounds coming from the maximum principle, we are able to control the collision process for all times (even for long times) and to optimize the strategy of Lanford [22] to get the following convergence.

Theorem 2.2. *Let φ_N^0 be a sequence of Lipschitz functions defined on $\mathbf{T}_\lambda^d\times\mathbf{R}^d$, satisfying* (2.10)*. Assume moreover that*

$$\mu_N\ll\sqrt{\log\log N} \quad\textit{and}\quad \Lambda_N\ll N^{\frac{2}{d^2-1}}$$

in the Boltzmann-Grad scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$.

Consider the initial distribution (2.9)*, describing the state of a tagged particle in a background of $N-1$ particles at equilibrium. Then the distribution $\lambda^d f_N^{(1)}$ of the tagged particle is asymptotic, as $N\to\infty$ under the Boltzmann-Grad scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv1$, to the solution g_N of the linear Boltzmann equation* (1.2) *with hard-sphere cross-section and with initial data*

$$g_N^0(z_1) := \lambda^d\left(\int_{\mathbf{T}_\lambda^d\times\mathbf{R}^d} M_\beta(v)\varphi_N^0(z)\,dz\right)^{-1} M_\beta(v_1)\varphi_N^0(z_1) \tag{2.11}$$

on any time interval $[0,t_N]$, where $t_N=o(\mu_N^{-1/2}\sqrt{\log\log N})$. More precisely, there is a constant C depending only on d and β such that

$$\left\|\lambda^d f_N^{(1)}-g_N\right\|_{L^\infty([0,t_N]\times\mathbf{T}_\lambda^d\times\mathbf{R}^d)} \leq C\mu_N^2\left(\frac{t_N^2}{\log\log N}\right)^2. \tag{2.12}$$

In [24], Lebowitz and Spohn derived the linear Boltzmann equation for any time $T > 0$ (independent of N). Compared to [24], our approach leads to quantitative estimates on the convergence up to times diverging when $N \to \infty$. As we shall see, this is the key to derive the diffusive limit in Theorem 2.3.

Remark 2.4. *The* $\log\log N$ *restriction on the time of convergence seems unavoidable by our strategy: as will be clear in the proof, on a time span t, the constant of continuity we find for the solution map is of the type t^J where J is the number of particles involved in the collision tree of the tagged particle; that number J will be shown to be of the order of C^t.*

As usual in this context, the precise convergence statement we shall establish concerns the BBGKY and Boltzmann hierarchies, and Theorem 2.2 is only a byproduct of this result considering the first marginal. The main difference with the usual strategy to prove convergence is that the symmetry is partially broken due to the fact that one particle is distinguished from the others. In other words $f_{N|t=0}$ is symmetric with respect to $z_2, \dots z_N$ but not to z_1, and this property is preserved by the dynamics. The quantities we shall consider are the following

$$f_N^{(s)}(t, Z_s) := \int f_N(t, Z_N) dz_{s+1} \dots dz_N$$

so $f_N^{(1)}$ is exactly the distribution of the tagged particle, and $f_N^{(s)}$ is the correlation between this tagged particle and $(s-1)$ particles of the background.

Note that the classical symmetry arguments used to establish the BBGKY hierarchy, i.e. the evolution equations for the marginals $f_N^{(s)}(t, Z_s)$, only deal with the particles we add by collisions to the sub-system Z_s under consideration. In particular, the equation in the BBGKY hierarchy will not be modified at all since - by convention - the tagged particle is numbered 1 and always belongs to the sub-system under consideration.

Remark 2.5. *Note that in the previous statement, the solution g_N to the linear Boltzmann equation depends on N via*

- *the domain $\mathbf{T}_\lambda$ since λ may depend on N,*
- *the initial data g_N^0.*

Since we shall study the convergence of the hierarchy, we introduce the following notation:

$$g_N^{(s)}(Z_s) := \left(\prod_{i=2}^{s} M_\beta(v_i) \right) g_N(z_1) \, ,$$

but we insist that $g_N^{(s)}$ are not defined as the marginals of some N-particle density. They are not even marginals of one another if $\lambda_N \neq 1$.

Remark 2.6. *We have chosen to deal with the tagged particle in a relatively symmetric way compared to other particles, having in mind to also consider the feedback of this particle on the background in a forthcoming work.*
In particular, the distribution of the tagged particle is assumed to be a fluctuation around the Maxwellian state $\lambda_N^{-d} M_\beta$, meaning that the local concentration of the tagged particle is almost homogeneous in the physical space, which is possible only because the domain has a finite (possibly large) volume λ_N^d. We therefore expect the quantity $\lambda_N^d f_N^{(1)}$ to have a finite limit as $N \to \infty$. If $\lambda_N \to \infty$, we get asymptotically the linear Boltzmann equation set in the whole space, but the normalization provides the following functional setting

$$M_\beta(\varphi^0 - 1) \in L^1(\mathbf{R}^d \times \mathbf{R}^d), \qquad \int M_\beta(\varphi^0 - 1) dx dv = 0 \, .$$

From a probabilistic point of view, it would be more natural to consider a tagged particle which is localized in space, typically with density

$$\varphi_N^0(x,v) = \mu_N \chi(\mu_N^{-1/d} x)\psi(v) \text{ where } \mu_N \to \infty. \tag{2.13}$$

Then the quantities which are expected to have a finite limit as $N \to \infty$ *are* $f_N^{(1)}$ *(without the normalization factor* λ_N^{-d}*), and more generally* $f_N^{(s)} \lambda_N^{d(s-1)}$*. Up to this disymmetrization between particle numbered 1 and other particles, all the arguments of the present paper can be reproduced identically. In particular, by this way, we would obtain the linear Boltzmann equation with Dirac mass as initial data even in the whole space.*

Theorem 2.2 proves that the linear Boltzmann equation is a good asymptotics of the hard-sphere dynamics, even for long times. It further provides a rather good estimate on the approximation error. Up to a suitable rescaling of space and time, we can therefore obtain diffusive limits. Let λ_N be the size of the torus which is now growing with N. In the macroscopic limit, the rescaled trajectory of the tagged particle is defined by

$$\Xi(\tau) = \frac{1}{\lambda_N} x_1\big(\lambda_N^2 \tau\big) \in \mathbf{T}^d, \tag{2.14}$$

where $\mathbf{T}^d$ has side length equal to 1. The distribution of $\Xi(\tau)$ is given by $\lambda_N^d f_N^{(1)}(\lambda_N^2 \tau, \lambda_N y, v)$. Let ζ be in $(0, \frac{1}{4d})$. We are going to consider initial data of the tagged particle which are localized in a mesoscopic neighborhood of the origin, i.e. in a very large domain in microscopic units but concentrated close to the origin after rescaling in the macroscopic domain $\mathbf{T}^d$. Let $\chi \geq 0$ be a smooth function defined in $\mathbf{R}^d$ supported by the unit ball and, such that $\int_{\mathbf{R}^d} \chi(y) dy = 1$. The initial data for the tagged particle are distributed according to

$$\forall x \in \mathbf{T}^d_{\lambda_N}, \qquad \rho_{N,0}(x) = \lambda_N^{d\zeta} \, \chi\left(\frac{x}{\lambda_N^{1-\zeta}}\right). \tag{2.15}$$

Note that this corresponds to the approximation of the Dirac mass (2.13) with $\mu_N = \lambda_N^{d(1-\zeta)}$, up to the normalization factor λ_N^{-d}.

Theorem 2.3. *Consider the dynamics* (2.1) *on* $\mathbf{T}^d_{\lambda_N} \times \mathbf{R}^d$ *of a tagged particle in a background of* $N-1$ *particles at equilibrium distributed initially according to* (2.9) *with* $\varphi_N^0(z_1) = \rho_{N,0}(x_1)$ *(see* (2.15)*) and* $\lambda_N = o((\log\log N)^{1/5})$*.*
Then the scaled distribution $\lambda_N^d f_N^{(1)}(\lambda_N^2 \tau, \lambda_N y, v)$ *of the tagged particle converges as* $N \to \infty$ *(in the Boltzmann-Grad scaling* $N\varepsilon^{d-1}\lambda_N^{-d} \equiv 1$*) in the sense of measures, i.e. in law, to the solution* $M_\beta \, \rho$ *of the linear heat equation*

$$\partial_\tau \rho - \kappa \Delta_y \rho = 0 \text{ on } \mathbf{R}^+ \times \mathbf{T}^d, \qquad \rho_{|\,\tau=0} = \delta_0\,, \tag{2.16}$$

where the diffusion coefficient κ *is given by* (6.10)*.*

More generally, the process Ξ *converges weakly towards a brownian motion with variance* κ*.*

Remark 2.7. *By considering a torus the size of which grows faster than* λ_N*, we obtain a dynamics which is defined asymptotically on the whole space*

$$\partial_\tau \rho - \kappa \Delta_y \rho = 0 \text{ on } \mathbf{R}^+ \times \mathbf{R}^d, \qquad \rho_{|t=0}(x) = \delta_0\,.$$

3. Uniform a priori estimates

As mentioned in the introduction, one of the crucial points in our proof of Theorem 2.2 is the fact that global in time a priori estimates are available. Indeed, even for stationary solutions, the estimates obtained by the usual Cauchy-Kowaleswki argument blow up after a short time: this is due to the fact that cancellations between the gain and loss terms of the collision operator are not taken into account.

In this section we obtain uniform a priori estimates for the marginals of the solution to the Liouville equation, thanks to the maximum principle. These global in time a priori estimates will be the key to prove that the collision process can be controlled, in the sense that collision trees with some "branching concentration" are very unlikely. Before doing so we discuss the asymptotic factorization of the initial data.

3.1. Asymptotic factorization of the initial data.

Let us prove the following result.

Proposition 3.1. *Let φ_N^0 be a sequence of continuous functions satisfying* (2.10). *Define the associated initial data f_N^0 as in* (2.9) *and the Boltzmann associated initial data as in* (2.11). *Given $\beta > 0$ there is a constant $C > 0$ such that for any fixed $s \in \mathbf{N}^*$, the marginal of order s*

$$f_N^{0(s)}(Z_s) := \int f_N^0(Z_N)\, dz_{s+1} \dots dz_N$$

satisfies, as $N \to \infty$ in the scaling $N\varepsilon^{d-1}\lambda^{-d} \equiv 1$,

$$\Big| \Big(\lambda^{ds} f_N^{0(s)} - g_N^{0(s)} \Big) \mathbf{1}_{\mathcal{D}_\varepsilon^s} \Big| \leq C\varepsilon \mu_N^2 s M_\beta^{\otimes s} \,,$$

where $g_N^{0(1)} = g_N^0$ is defined in (2.11) *and*

$$g_N^{0(s)}(Z_s) := \Big(\prod_{i=2}^s M_\beta(v_i) \Big)\, g_N^0(z_1)\,. \tag{3.1}$$

Proof. The proof follows exactly the same lines as the proof of Proposition 2.1 given in Appendix A. We recall indeed that thanks to (A.1) we have

$$1 \leq \lambda^{ds} \bar{\mathcal{Z}}_N^{-1} \bar{\mathcal{Z}}_{N-s} \leq (1 - \varepsilon\kappa_d)^{-s}\,, \tag{3.2}$$

as $N \to \infty$ under the scaling $N\varepsilon^{d-1}\lambda^{-d} \equiv 1$, where we recall that

$$\bar{\mathcal{Z}}_N := \int_{\mathbf{T}_\lambda^{dN} \times \mathbf{R}^{dN}} \Big(\frac{\beta}{2\pi} \Big)^{\frac{dN}{2}} \exp(-\beta H_N(Z_N)) \mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N)\, dZ_N\,.$$

Next we claim that

$$\bar{\mathcal{Z}}_N^{-1} \mathcal{Z}_N \sim \lambda^{-d} \int_{\mathbf{T}_\lambda^d \times \mathbf{R}^d} M_\beta(v) \varphi_N^0(z)\, dz\,. \tag{3.3}$$

Assuming that result to be true, the computations of Appendix A, Second step, become

$$\begin{aligned} f_N^{0(s)}(Z_s) = \mathcal{Z}_N^{-1} \mathbf{1}_{Z_s \in \mathcal{D}_\varepsilon^s} \Big(\frac{\beta}{2\pi} \Big)^{\frac{sd}{2}} \exp\Big(-\frac{\beta}{2} |V_s|^2 \Big) \varphi_N^0(z_1) \\ \times \int_{\mathbf{T}_\lambda^{d(N-s)}} \prod_{s+1 \leq i \neq j \leq N} \mathbf{1}_{|x_i - x_j| > \varepsilon} \prod_{i' \leq s < j'} \mathbf{1}_{|x_{i'} - x_{j'}| > \varepsilon}\, dX_{(s+1,N)}\,, \end{aligned}$$

hence

$$f_N^{0(s)}(Z_s) = \mathcal{Z}_N^{-1} \mathbf{1}_{Z_s \in \mathcal{D}_\varepsilon^s} M_\beta^{\otimes s} \Big(\bar{\mathcal{Z}}_{N-s} - \bar{\mathcal{Z}}^\flat_{(s+1,N)} \Big) \varphi_N^0(z_1)$$

with the notation

$$\bar{\mathcal{Z}}^\flat_{(s+1,N)} := \int_{\mathbf{T}_\lambda^{d(N-s)}} \Big(1 - \prod_{i\leq s<j} \mathbf{1}_{|x_i-x_j|>\varepsilon}\Big) \prod_{s+1\leq i\neq j\leq N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\, dX_{(s+1,N)}\,.$$

The result then follows from (3.2) and (3.3) along with the following estimate, proved in Appendix A:

$$\lambda^{ds}\bar{\mathcal{Z}}_N^{-1}\bar{\mathcal{Z}}^\flat_{(s+1,N)} \leq \varepsilon s\kappa_d\big(1-\varepsilon\kappa_d\big)^{-(s+1)}\,.$$

We conclude exactly as in the proof of Proposition 2.1.

So let us prove (3.3). We have

$$\begin{aligned}\mathcal{Z}_N &= \int M_\beta(v_1)\varphi^0_N(z_1) \prod_{1\leq i\neq j\leq N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\, dX_N dv_1\\ &= \int M_\beta(v_1)\varphi^0_N(z_1)\big(\prod_{2\leq i\neq j\leq N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\big)\Big(1+\big(\prod_{2\leq i\leq N}\mathbf{1}_{|x_i-x_1|>\varepsilon}-1\big)\Big)\, dX_N dv_1\,.\end{aligned}$$

On the one hand we have

$$\int M_\beta(v_1)\varphi^0_N(z_1)\big(\prod_{2\leq i\neq j\leq N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\big)\, dX_N dv_1 = \int M_\beta(v_1)\varphi^0_N(z_1)\, dz_1\, \bar{\mathcal{Z}}_{N-1}$$

and on the other hand

$$\Big|\int M_\beta(v_1)\varphi^0_N(z_1)\big(\prod_{2\leq i\neq j\leq N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\big)\big(\prod_{2\leq i\leq N}\mathbf{1}_{|x_i-x_1|>\varepsilon}-1\big)\, dX_N dv_1\Big| \leq N\varepsilon^d\mu_N\bar{\mathcal{Z}}_{N-1}$$

so we find, thanks to (3.2),

$$\begin{aligned}\mathcal{Z}_N &= \int M_\beta(v_1)\varphi^0_N(z_1)\, dz_1\, \bar{\mathcal{Z}}_{N-1} + O(N\varepsilon^d\mu_N\bar{\mathcal{Z}}_{N-1})\\ &\sim \lambda^{-d}\bar{\mathcal{Z}}_N \int M_\beta(v_1)\varphi^0_N(z_1)\, dz_1 + O(\varepsilon\mu_N\bar{\mathcal{Z}}_N)\end{aligned}$$

whence the result, recalling that $\mu_N \ll (\log\log N)^{\frac12}$. □

3.2. **A priori estimates coming from the maximum principle.** The major advantage in considering such initial data as (2.9) is that we get very easily uniform a priori bounds, using only the maximum principle for the Liouville equation (2.4).

Proposition 3.2. *Let φ^0_N be a continuous function satisfying* (2.10)*. For any fixed N, denote by f_N the solution to the Liouville equation* (2.4) *with initial data* (2.9)*, and by $f^{(s)}_N$ its marginal of order s:*

$$f^{(s)}_N(t,Z_s) := \int f_N(t,Z_N)\, dz_{s+1}\dots dz_N\,. \tag{3.4}$$

Then, for any $s\geq 1$, we have the following uniform bound (with respect to time):

$$\sup_t f^{(s)}_N(t,Z_s) \leq \mu_N^2 M^{(s)}_{N,\beta}(Z_s) \leq \lambda^{-ds}\mu_N^2\big(1-\varepsilon\kappa_d\big)^{-s} M_\beta^{\otimes s}(V_s)\,, \tag{3.5}$$

where $M^{(s)}_{N,\beta}$ is the marginal of order s of the Gibbs measure $M_{N,\beta}$, defined by (2.7)*, and κ_d denotes the volume of the unit ball in $\mathbf{R}^d$.*

Note here that although the variable z_1 does not play at all a symmetric role with respect to $z_2,\dots z_N$, the upper bound (3.5) does not see this asymmetry.

Proof. We first remark that

$$\mathcal{Z}_N = \int_{\mathbf{T}_\lambda^{dN}\times\mathbf{R}^{dN}} \left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(Z_N))\varphi_N^0(z_1)\, dZ_N$$
$$\geq \frac{1}{\mu_N}\int_{\mathbf{T}_\lambda^{dN}\times\mathbf{R}^{dN}} \left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(Z_N)) dZ_N \geq \frac{1}{\mu_N}\bar{\mathcal{Z}}_N$$

thanks to (2.6).
Combining this estimate with the L^∞-bound on φ_N^0, we get

$$f_N^0(Z_N) \leq \mu_N^2 M_{N,\beta}(Z_N)\,.$$

Since the maximum principle holds for the Liouville equation (2.4), and as the Gibbs measure $M_{N,\beta}$ is a stationary solution of (2.4), we get for all $t\in\mathbf{R}$

$$f_N(t,Z_N) \leq \mu_N^2 M_{N,\beta}(Z_N)\,.$$

The inequalities for the marginals follow by integration. We write

$$M_{N,\beta}^{(s)}(Z_s) = \frac{1}{\bar{\mathcal{Z}}_N}\int_{\mathbf{T}_\lambda^{dN}\times\mathbf{R}^{dN}} \left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(Z_N))\mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N)\, dz_{s+1}\dots dz_N$$
$$\leq \bar{\mathcal{Z}}_N^{-1} M_\beta^{\otimes s}(Z_s)\int_{\mathbf{T}_\lambda^{d(N-s)}} \prod_{s+1\leq i\neq j\leq N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\, dz_{s+1}\dots dz_N$$
$$\leq \bar{\mathcal{Z}}_N^{-1}\bar{\mathcal{Z}}_{N-s} M_\beta^{\otimes s}(Z_s)\,,$$

and the result follows from (A.1). □

On the other hand, it is easy to check that the maximum principle holds for the linear Boltzmann equation (which can be seen as a Markov process), so the following result holds.

Proposition 3.3. *Let φ_N^0 be a continuous function satisfying* (2.10)*. For any fixed N, denote by g_N the solution to the linear Boltzmann equation* (1.2) *with hard-sphere cross-section and with initial data*

$$g_N^0(z_1) := \lambda^d \left(\int_{\mathbf{T}_\lambda^d\times\mathbf{R}^d} M_\beta(v)\varphi_N^0(z)\, dz\right)^{-1} M_\beta(v_1)\varphi_N^0(z_1)$$

and we define

$$g_N^{(s)}(t,Z_s) := \left(\prod_{i=2}^s M_\beta(v_i)\right) g_N(t,z_1)\,.$$

Then, for any $s\geq 1$, we have the following uniform bound (with respect to time):

$$\sup_t g_N^{(s)}(t,Z_s) \leq \mu_N^2 M_\beta^{\otimes s}(V_s)\,. \tag{3.6}$$

The uniform a priori bounds (3.5)-(3.6) will replace the analytical-type estimates obtained by Cauchy-Kowalewski arguments in the proof of Lanford ([22, 9]). The important point is that

- they will enable us to use dominated convergence arguments to get rid of high order correlations (due to “big collision trees”);
- they are propagated globally in time, so the usual short time estimates can be iterated to produce a global convergence result.

4. Control of the branching process

Lanford's convergence proof on a short time τ (determined by the weighted norm of the initial data) is based on a truncation of "pathological" collision trees, defined by a too large number of branches created in the time interval $[0,\tau]$ (typically greater than $n_\varepsilon = O(|\log\varepsilon|)$, see [14] for a quantitative estimate of the truncation parameter). Here the global bound coming from the maximum principle will enable us to iterate this truncation process on any time interval of size τ.

4.1. **The series expansion.**

In the case of hard-spheres, a straightforward computation based on Green's formula (see [14] for a rigorous justification using weak formulations) leads to the following the following **BBGKY hierarchy** for $s < N$:

$$(\partial_t + \sum_{i=1}^s v_i \cdot \nabla_{x_i}) f_N^{(s)}(t, Z_s) = \big(C_{s,s+1} f_N^{(s+1)}\big)(t, Z_s) \tag{4.1}$$

on $\mathcal{D}_\varepsilon^s$, where the collision term is defined by

$$\begin{aligned}&\big(C_{s,s+1} f_N^{(s+1)}\big)(Z_s)\\ &:= (N-s)\varepsilon^{d-1} \sum_{i=1}^s \int_{\mathbf{S}^{d-1}\times\mathbf{R}^d} f_N^{(s+1)}(\dots, x_i, v_i^*, \dots, x_i + \varepsilon\nu, v_{s+1}^*)\Big((v_{s+1} - v_i)\cdot\nu\Big)_+ d\nu dv_{s+1}\\ &- (N-s)\varepsilon^{d-1} \sum_{i=1}^s \int_{\mathbf{S}^{d-1}\times\mathbf{R}^d} f_N^{(s+1)}(\dots, x_i, v_i, \dots, x_i + \varepsilon\nu, v_{s+1})\Big((v_{s+1} - v_i)\cdot\nu\Big)_- d\nu dv_{s+1}\end{aligned} \tag{4.2}$$

where $\mathbf{S}^{d-1}$ denotes the unit sphere in $\mathbf{R}^d$.

Note that the collision integral is split into two terms according to the sign of $(v_i - v_{s+1})\cdot\nu$ and used the trace condition on $\partial\mathcal{D}_\varepsilon^N$

$$f_N(t, \dots, x_i, v_i, \dots, x_i + \varepsilon\nu, v_{s+1}, \dots) = f_N(t, \dots, x_i, v_i^*, \dots, x_i + \varepsilon\nu, v_{s+1}^*, \dots)$$

where

$$v_i^* := v_i - (v_i - v_{s+1})\cdot\nu\,\nu, \qquad v_{s+1}^* := v_{s+1} + (v_i - v_{s+1})\cdot\nu\,\nu\,,$$

to express all quantities in terms of pre-collisional configurations. The closure for $s = N$ is given by the Liouville equation (2.4).

Mild solutions of the BBGKY hierarchy are thus defined by Duhamel's formula

$$f_N^{(s)}(t) = \mathbf{S}_s(t) f_N^{(s)}(0) + \int_0^t \mathbf{S}_s(t - t_1) C_{s,s+1} f_N^{(s+1)}(t_1)\, dt_1\,, \tag{4.3}$$

denoting by $\mathbf{S}_s$ the semi-group associated to free transport in $\mathcal{D}_\varepsilon^s$ with specular reflection on the boundary.

Given the special role played by the initial data (which is the reference to determine the notion of pre-collisional and post-collisional configurations), it is then natural to express solutions of the BBGKY hierarchy in terms of a series of operators applied to the initial marginals. The starting point in Lanford's proof is therefore the **iterated Duhamel formula**

$$\begin{aligned}f_N^{(s)}(t) = \sum_{n=0}^{N-s} \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} &\mathbf{S}_s(t - t_1) C_{s,s+1} \mathbf{S}_{s+1}(t_1 - t_2) C_{s+1,s+2}\\ &\dots \mathbf{S}_{s+n}(t_n) f_N^{(s+n)}(0)\, dt_n \dots dt_1\,.\end{aligned}$$

To simplify notations, we define the operators $Q_{s,s}(t) = \mathbf{S}_s(t)$ and for $n \geq 1$

$$Q_{s,s+n}(t) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} \mathbf{S}_s(t-t_1) C_{s,s+1} \mathbf{S}_{s+1}(t_1-t_2) C_{s+1,s+2} \dots \mathbf{S}_{s+n}(t_n)\, dt_n \dots dt_1 \tag{4.4}$$

so that

$$f_N^{(s)}(t) = \sum_{n=0}^{N-s} Q_{s,s+n}(t) f_N^{(s+n)}(0)\,. \tag{4.5}$$

To obtain the Boltzmann series expansion we start with the expansion (4.5) and compute the formal limit of the collision operator $\lambda^{-dn} Q_{s,s+n}$. This is classically given by (recall that $(N-s)\varepsilon^{d-1}\lambda^{-d} \sim 1$)

$$Q^0_{s,s+n}(t) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} \mathbf{S}^0_s(t-t_1) C^0_{s,s+1} \mathbf{S}^0_{s+1}(t_1-t_2) C^0_{s+1,s+2} \dots \mathbf{S}^0_{s+n}(t_n)\, dt_n \dots dt_1 \tag{4.6}$$

where $\mathbf{S}^0_s$ denotes the free flow of s particles on $\mathbf{T}^{ds}_\lambda \times \mathbf{R}^{ds}$, and $C^0_{s,s+1}$ are the limit collision operators defined by

$$\begin{aligned}
\big(C^0_{s,s+1} g^{(s+1)}\big)(Z_s) = &\sum_{i=1}^s \int g^{(s+1)}(\dots, x_i, v_i^*, \dots, x_i, v_{s+1}^*) \Big((v_{s+1}-v_i)\cdot\nu \Big)_+ d\nu dv_{s+1}\,, \\
&- \sum_{i=1}^s \int g^{(s+1)}(\dots, x_i, v_i, \dots, x_i, v_{s+1}) \Big((v_{s+1}-v_i)\cdot\nu \Big)_- d\nu dv_{s+1}\,.
\end{aligned} \tag{4.7}$$

Then the iterated Duhamel formula for the Boltzmann equation takes the form

$$g_N^{(s)}(t) = \sum_{n\geq 0} Q^0_{s,s+n}(t) g_N^{(s+n)}(0)\,. \tag{4.8}$$

Notice that in the special case when the initial data to the Boltzmann hierarchy is given by (3.1), then

$$g_N^{(s)}(t, Z_s) = \left(\prod_{i=2}^s M_\beta(v_i) \right) g_N(t, z_1)\,, \tag{4.9}$$

solves the hierarchy, as soon as g_N satisfies the linear Boltzmann equation (1.2).

4.2. **Continuity estimates for the collision operators.** To get uniform estimates with respect to N, the usual strategy is to use some Cauchy-Kowalewski argument, over-estimating all collision integrals by their absolute value. In the following we therefore denote by $|Q|_{s,s+n}$ the operator obtained by summing the absolute values of all elementary contributions, and similarly for $|Q^0|_{s,s+n}$.

We define $X_{\varepsilon,k,\alpha}$ the space of continuous functions f_k defined on $\mathcal{D}^k_\varepsilon$ such that

$$\|f_k\|_{\varepsilon,k,\alpha} := \sup_{Z_k \in \mathcal{D}^k_\varepsilon} \Big| f_k(Z_k)\ \exp\big(\alpha H_k(Z_k)\big) \Big| < \infty\,,$$

and similarly $X_{0,k,\alpha}$ is the space of continuous functions g_k defined on $\mathbf{T}^{dk}_\lambda \times \mathbf{R}^{dk}$ such that

$$\|g_k\|_{0,k,\alpha} := \sup_{Z_k \in \mathbf{T}^{dk}_\lambda \times \mathbf{R}^{dk}} \Big| g(Z_k) \exp\,\big(\alpha H_k(Z_k)\big) \Big| < \infty\,.$$

Lemma 4.1. *There is a constant C_d depending only on d such that for all $s,n\in\mathbf{N}$, $t\geq 0$, the operators $|Q|_{s,s+n}(t)$ and $|Q^0|_{s,s+n}(t)$ satisfy the following continuity estimates: for all f_{s+n} in $X_{\varepsilon,s+n,\alpha}$, $|Q|_{s,s+n}(t)f_{s+n}$ belongs to $X_{\varepsilon,s,\frac{\alpha}{2}}$ and*

$$\Big\| |Q|_{s,s+n}(t)f_{s+n}\Big\|_{\varepsilon,s,\frac{\alpha}{2}} \leq e^{s-1}\left(\frac{C_d\lambda^d t}{\alpha^{\frac{d+1}{2}}}\right)^n \|f_{s+n}\|_{\varepsilon,s+n,\alpha}\,. \tag{4.10}$$

Similarly for all g_{s+n} in $X_{0,s+n,\alpha}$, $|Q^0|_{s,s+n}(t)g_{s+n}$ belongs to $X_{0,s,\frac{\alpha}{2}}$ and

$$\Big\| |Q^0|_{s,s+n}(t)g_{s+n}\Big\|_{0,s,\frac{\alpha}{2}} \leq e^{s-1}\left(\frac{C_d t}{\alpha^{\frac{d+1}{2}}}\right)^n \|g_{s+n}\|_{0,s+n,\alpha}\,. \tag{4.11}$$

Note that in the case of tensorized initial data such as (2.9), the estimates coming from the maximum principle provide a control, which is global in time, on the weighted norms $\lambda^{dk}\|f_N^{(k)}(t)\|_{\varepsilon,k,\beta}$ for all $t\geq 0$. Indeed we have thanks to Proposition 3.2

$$\begin{aligned}\|f_N^{(k)}(t)\|_{\varepsilon,k,\beta} &= \sup_{Z_k\in\mathcal{D}_\varepsilon^k}\Big|f_N^{(k)}(t,Z_k)\ \exp\big(\beta H_k(Z_k)\big)\Big|\\ &\leq \mu_N^2\sup_{Z_k\in\mathcal{D}_\varepsilon^k} M_{N,\beta}^{(k)}(Z_k)\exp\big(\beta H_k(Z_k)\big)\\ &\leq \mu_N^2\lambda^{-dk}\big(1-\varepsilon\kappa_d\big)^{-k}\sup_{Z_k\in\mathcal{D}_\varepsilon^k} M_\beta^{\otimes k}(V_k)\exp\big(\beta H_k(Z_k)\big)\,.\end{aligned}$$

Thus for all $t\in\mathbf{R}$,

$$\lambda^{dk}\|f_N^{(k)}(t)\|_{\varepsilon,k,\beta}\leq \mu_N^2\big(1-\varepsilon\kappa_d\big)^{-k}\Big(\frac{\beta}{2\pi}\Big)^{kd/2}\,. \tag{4.12}$$

Similarly the initial data for the Boltzmann hierarchy defined in (3.1) satisfies

$$\|g_N^{0(k)}\|_{0,k,\beta}\leq \mu_N^2\Big(\frac{\beta}{2\pi}\Big)^{kd/2}\,, \tag{4.13}$$

and, by Proposition 3.3,

$$\|g_N^{(k)}(t)\|_{0,k,\beta}\leq \mu_N^2\Big(\frac{\beta}{2\pi}\Big)^{kd/2}\,. \tag{4.14}$$

Proof of Lemma 4.1. Estimate (4.10) is simply obtained from the fact that the transport operators preserve the weighted norms, along with the continuity of the elementary collision operators. We indeed recall from [14, Chapter 5] the following statements:

- the transport operators satisfy the identities
$$\begin{aligned}\|\mathbf{S}_k(t)f_k\|_{\varepsilon,k,\alpha}&=\|f_k\|_{\varepsilon,k,\alpha}\\ \|\mathbf{S}_k^0(t)g_k\|_{0,k,\alpha}&=\|g_k\|_{0,k,\alpha}\,.\end{aligned}$$
- the collision operators satisfy the following bounds in the Boltzmann-Grad scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$ (see Propositions 5.3.1 and 5.3.2 of [14]):
$$\begin{aligned}\big|C_{k,k+1}\ f_{k+1}(Z_k)\big| &\leq \lambda^d C_d\,\alpha^{-\frac{d}{2}}\Big(k\alpha^{-\frac12}+\sum_{1\leq i\leq k}|v_i|\Big)\exp\left(-\alpha H_k(Z_k)\right)\|f_{k+1}\|_{\varepsilon,k+1,\alpha}\\ \big|C^0_{k,k+1}\ g_{k+1}(Z_k)\big| &\leq C_d\,\alpha^{-\frac{d}{2}}\Big(k\alpha^{-\frac12}+\sum_{1\leq i\leq k}|v_i|\Big)\exp\left(-\alpha H_k(Z_k)\right)\|g_{k+1}\|_{0,k+1,\alpha}\,,\end{aligned} \tag{4.15}$$

  for some $C_d>0$ depending only on d.

The result then follows from piling together those inequalities (dispatching the exponential weight evenly on each occurence of a collision term). We notice that by Cauchy-Schwarz,

$$\sum_{1\leq i\leq k}|v_i|\exp\Big(-\frac{\alpha}{4n}\sum_{1\leq j\leq k}|v_j|^2\Big)\leq\Big(k\frac{2n}{\alpha}\Big)^{\frac12}\left(\sum_{1\leq i\leq k}\frac{\alpha}{2n}|v_i|^2\exp\Big(-\frac{\alpha}{2n}\sum_{1\leq j\leq k}|v_j|^2\Big)\right)^{1/2}$$
$$\leq\Big(\frac{2nk}{e\alpha}\Big)^{1/2}\leq\sqrt{\frac{2}{e\alpha}}(s+n-1)\,.$$

Each collision operator gives therefore a loss of $C\lambda^d\alpha^{-(d+1)/2}(s+n-1)$ together with a loss on the exponential weight, while the integration with respect to time provides a factor $t^n/n!$. By Stirling's formula, we have

$$\frac{(s+n-1)^n}{n!}\leq\exp\left(n\log\frac{n+s-1}{n}+n\right)\leq\exp(s-1+n)\,.$$

That proves the first statement in the lemma. The same arguments give the counterpart for the Boltzmann collision operator. □

The proof of Lanford's convergence result then relies on two steps:

(i) a short time bound for the series expansion expressing the correlations of the system of N particles and a similar bound for the corresponding quantities associated with the Boltzmann hierarchy;
(ii) the termwise convergence.

However after a short time (depending on the initial data), the series becomes divergent and the question of the convergence after that time is still open. The divergence of the series is partially due to the fact that possible cancellations are completely neglected in this strategy. Here we shall take advantage of the control by stationary solutions (the existence of which is obviously related to these cancellations) given by Proposition 3.2 to obtain a lifespan which does not depend on the initial data. Moreover we shall use a truncated series expansion instead of (4.5) and (4.8).

4.3. **Good collision trees.** From now on we assume that the BBGKY initial data takes the form (2.9) and the Boltzmann initial data takes the form (3.1).

Let us fix a (small) parameter $\tau>0$ and a sequence $\{n_k\}_{k\geq1}$ of integers to be tuned later. We shall study the dynamics up to time $t:=K\tau$ for some large integer K, by splitting the time interval $[0,t]$ into K intervals $\bigcup_{1\leq k\leq K}[(k-1)\tau,k\tau]$, and controling the number of collisions on each interval. In order to discard trajectories with a large number of collisions in the iterated Duhamel formula (4.5), we define "good" collision trees by the condition that they have less than n_k branch points on the interval $[t-k\tau,t-(k-1)\tau]$. Note that by construction, the trees are actually followed "backwards", from time t (large) to time 0.
As we are interested only in the asymptotic behaviour of the first marginal, we start by using (4.3) with $s=1$, during the time interval $[t-\tau,t]$: iterating Duhamel's formula n_1-1 times instead of $N-1$ as in (4.5), we have

$$f_N^{(1)}(t)=\sum_{j_1=1}^{n_1-1}Q_{1,j_1}(\tau)f_N^{(j_1)}(t-\tau)+R_{1,n_1}(t-\tau,t)\,,\tag{4.16}$$

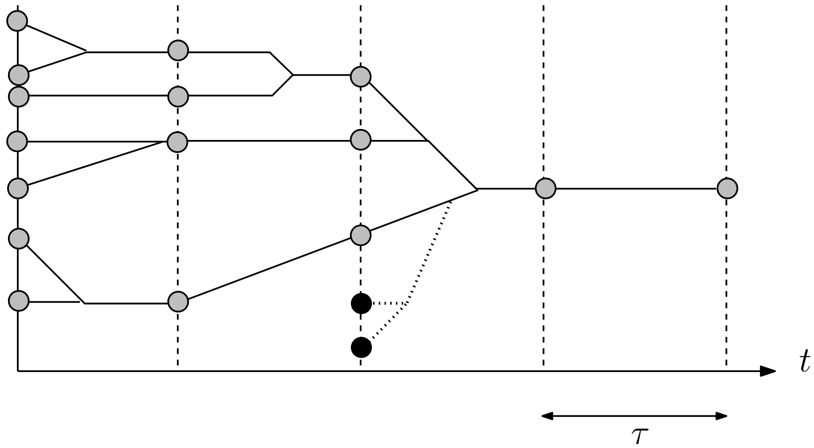

FIGURE 1. The tree including the two extra collisions in dotted lines occurring during $[t-2\tau, t-\tau]$ is not a good collision tree and in our procedure, it would be truncated at time $t-2\tau$. The tree without the dotted lines is a good collision tree with $t=4\tau$: the number of collisions during the k^{th}-time interval is less than $n_k - 1 = 2^k - 1$.

where

$$R_{1,n}(t',t) := \int_{t'}^{t}\int_{t'}^{t_1}\dots\int_{t'}^{t_{n-1}} \mathbf{S}_1(t-t_1)C_{1,2}\mathbf{S}_2(t_1-t_2)C_{2,3}\dots \\ \dots C_{n,n+1}f_N^{(n+1)}(t_n)\,dt_n\dots dt_1\,.$$

More generally we define $R_{k,n}$ as follows:

$$R_{k,n}(t',t) := \int_{t'}^{t}\int_{t'}^{t_1}\dots\int_{t'}^{t_{n-1}} \mathbf{S}_k(t-t_1)C_{k,k+1}\mathbf{S}_{k+1}(t_1-t_2)C_{k+1,k+2}\dots \\ \dots C_{k+n-1,k+n}f_N^{(k+n)}(t_n)\,dt_n\dots dt_1\,. \tag{4.17}$$

The term $R_{k,n}(t',t)$ accounts for trajectories originating at k points, and involving at least n collisions during the time-span $t-t'$. The idea is that if n is large then such a behaviour should be atypical and $R_{k,n}(t',t)$ should be negligible.

The first term on the right-hand side of (4.16) can be broken up again by iterating the Duhamel formula on the time interval $[t-2\tau, t-\tau]$ and truncating the contributions with more than n_2 collisions: this gives

$$f_N^{(1)}(t) = \sum_{j_1=1}^{n_1-1}\sum_{j_2=0}^{n_2-1} Q_{1,j_1}(\tau)Q_{j_1,j_1+j_2}(\tau)\,f_N^{(j_1+j_2)}(t-2\tau) \\ + R_{1,n_1}(t-\tau,t) + \sum_{j_1=1}^{n_1-1} Q_{1,j_1}(\tau)R_{j_1,n_2}(t-2\tau,t-\tau)\,.$$

Iterating this procedure K times and truncating the trajectories with at least n_k collisions during the time interval $[t-k\tau, t-(k-1)\tau]$, leads to the following expansion

$$f_N^{(1)}(t) = f_N^{(1,K)}(t) + R_N^K(t)\,, \tag{4.18}$$

where denoting $J_k := j_1 + \cdots + j_k$,

$$
\begin{aligned}
f_N^{(1,K)}(t) &:= \sum_{j_1=1}^{n_1-1} \dots \sum_{j_K=0}^{n_K-1} Q_{1,J_1}(\tau) Q_{J_1,J_2}(\tau) \dots Q_{J_{K-1},J_K}(\tau)\, f_N^{0(J_K)} \\
R_N^K(t) &:= \sum_{k=1}^{K} \sum_{j_1=1}^{n_1-1} \dots \sum_{j_{k-1}=0}^{n_{k-1}-1} Q_{1,J_1}(\tau) \dots Q_{J_{k-2},J_{k-1}}(\tau)\, R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau)\,.
\end{aligned}
\tag{4.19}
$$

By an appropriate choice of the sequence $\{n_k\}$, we are going to show that the main contribution to the density $f_N^{(1)}(t)$ is given by the first term in the right-hand side of (4.18) and that the second term vanishes asymptotically.

Next as in (4.19) we can write a truncated expansion for g_N (see (4.9)) as follows:

$$
g_N(t) = g_N^{(1,K)}(t) + R_N^{0,K}(t)\,, \tag{4.20}
$$

where with notation (3.1) and (4.9),

$$
\begin{aligned}
g_N^{(1,K)}(t) &:= \sum_{j_1=1}^{n_1-1} \dots \sum_{j_K=0}^{n_K-1} Q^0_{1,J_1}(\tau) Q^0_{J_1,J_2}(\tau) \dots Q^0_{J_{K-1},J_K}(\tau)\, g_N^{0(J_K)} \\
R_N^{0,K}(t) &:= \sum_{k=1}^{K} \sum_{j_1=1}^{n_1-1} \dots \sum_{j_{k-1}=0}^{n_{k-1}-1} Q^0_{1,J_1}(\tau) \dots Q^0_{J_{k-2},J_{k-1}}(\tau)\, R^0_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau)
\end{aligned}
\tag{4.21}
$$

with

$$
\begin{aligned}
R^0_{k,n}(t',t) := \int_{t'}^{t} \int_{t'}^{t_1} \dots \int_{t'}^{t_{n-1}} \mathbf{S}^0_k(t-t_1) C^0_{k,k+1} \mathbf{S}^0_{k+1}(t_1-t_2) C^0_{k+1,k+2} \dots \\
\dots C^0_{k+n-1,k+n} g_N^{(k+n)}(t_n)\, dt_n \dots dt_1\,.
\end{aligned}
\tag{4.22}
$$

4.4. **Estimates of the remainders.** Recall that we have the global uniform bounds (4.12) and (4.14):

$$
\begin{aligned}
\lambda^{dk} \| f_N^{(k)}(t) \|_{\varepsilon,k,\beta} &\leq \mu_N^2 \big(1 - \varepsilon \kappa_d\big)^{-k} \Big(\frac{\beta}{2\pi}\Big)^{kd/2} \\
\| g_N^{(k)}(t) \|_{0,k,\beta} &\leq \mu_N^2 \Big(\frac{\beta}{2\pi}\Big)^{kd/2} .
\end{aligned}
$$

Since we expect the particles to undergo in average one collision per unit of time, the growth of collision trees is typically exponential (as for a random tree). Pathological trees are therefore those with super exponential growth. There are actually two natural ways of defining such pathological trees

- either by choosing some fixed τ (given for instance by Lanford's proof) and $\log n_k \gg k$;
- or by fixing $n_k = 2^k$ and letting the elementary time interval $\tau \to 0$.

We will choose the latter option.

Proposition 4.2. *Under the assumptions of Theorem 2.2, the following holds. Let $n_k = 2^k$, for $k \geq 1$. Then there exist $c, C, \gamma_0 > 0$ depending only on d and β such that for any $t > 0$ and any $\gamma \leq \gamma_0$, choosing*

$$
\tau \leq \frac{c\gamma}{t} \quad \textit{and} \quad K = t/\tau \ \textit{integer} \tag{4.23}
$$

we get

$$
\lambda^d \left\| R_N^K(t) \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} + \left\| R_N^{0,K}(t) \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} \leq C \mu_N^2 \gamma^2\,. \tag{4.24}
$$

Proof. We are going to bound

$$\left\| Q_{1,J_1}(\tau) \dots Q_{J_{k-2},J_{k-1}}(\tau)\, R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau) \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)}$$

for each term in the remainder R_N^K. The exact distribution of collisions in the last $k-1$ intervals is not needed and it is enough to estimate directly

$$\left\| |Q|_{1,J_{k-1}}((k-1)\tau)\, R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau) \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} .$$

Applying Lemma 4.1, one has (denoting generically by C_d any constant depending only on d)

$$\begin{aligned} &\left\| |Q|_{1,J_{k-1}}((k-1)\tau)\, R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau) \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} \\ &\quad \leq \left(\frac{C_d \lambda^d \,(k-1)\tau}{\beta^{(d+1)/2}} \right)^{J_{k-1}-1} \left\| R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau) \right\|_{\varepsilon, J_{k-1}, \beta/2} . \end{aligned}$$

Then arguing as in the proof of Lemma 4.1, one can write

$$\begin{aligned} &\left\| |Q|_{1,J_{k-1}}((k-1)\tau)\, R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau) \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} \\ &\quad \leq \left(\frac{C_d \lambda^d (k-1)\tau}{\beta^{(d+1)/2}} \right)^{J_{k-1}-1} \left(\frac{C_d \lambda^d \tau}{\beta^{(d+1)/2}} \right)^{n_k} \sup_{t \geq 0} \| f_N^{(J_{k-1}+n_k)}(t) \|_{\varepsilon, J_{k-1}+n_k, \beta} \\ &\quad \leq \mu_N^2 \lambda^{-d} \beta^{\frac{d}{2}} \left(\frac{C_d}{\sqrt{\beta}} \right)^{J_{k-1}+n_k-1} t^{J_{k-1}-1}\, \tau^{n_k} , \end{aligned} \tag{4.25}$$

thanks to (4.12) and recalling that $(k-1)\tau \leq t$.
Note that $\mathcal{N}_j := n_1 + \dots + n_j = 2^{j+1} - 2 \leq 2n_j$. Then, since $J_{k-1} \leq \mathcal{N}_{k-1}$, one has

$$\begin{aligned} &\lambda^d \left\| |Q|_{1,J_{k-1}}((k-1)\tau)\, R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau) \ \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} \\ &\quad \leq \mu_N^2 \beta^{d/2} \exp\left(2^k \Big(\log \frac{C_d^2}{\beta} + \log t + \log \tau \Big) \right) . \end{aligned}$$

Therefore, choosing

$$\tau \leq \left(\frac{\beta}{C_d^2} \right) \frac{\gamma}{t} ,$$

one has

$$\lambda^d \left\| |Q|_{1,J_{k-1}}((k-1)\tau)\, R_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau) \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} \leq \mu_N^2 \beta^{d/2} \exp\left(2^k \log \gamma \right) . \tag{4.26}$$

This implies

$$\begin{aligned} &\lambda^d \left\| R_N^K \right\|_{L^\infty(\mathbf{T}^d_\lambda \times \mathbf{R}^d)} \\ &\leq \mu_N^2 \beta^{d/2} \sum_{k=1}^K \Big(\prod_{i=1}^k n_i \Big) \ \exp\left(2^k \log \gamma \right) \leq \mu_N^2 \beta^{d/2} \sum_{k=1}^K \ \exp\left(k(k+1) + 2^k \log \gamma \right) \\ &\leq C_d \mu_N^2 \beta^{d/2} \sum_{k=1}^K \ \exp\left(2k \log \gamma \right) . \end{aligned}$$

Thus we get (4.24). The argument is identical in the case of the Boltzmann hierarchy: estimate (6.12) becomes

$$\begin{aligned}\Big\| |Q|^0_{1,J_{k-1}}((k-1)\tau)\, R^0_{J_{k-1},n_k}(t-k\tau, t-(k-1)\tau)\Big\|_{L^\infty(\mathbf{T}^d_\lambda\times\mathbf{R}^d)} \\ \leq \Big(\frac{C_d(k-1)\tau}{\beta^{(d+1)/2}}\Big)^{J_{k-1}-1}\Big(\frac{C_d\tau}{\beta^{(d+1)/2}}\Big)^{n_k}\sup_{t\geq 0}\|g_N^{(J_{k-1}+n_k)}(t)\|_{0,J_{k-1}+n_k,\beta} \\ \leq \mu_N^2\beta^{\frac d2}\Big(\frac{C_d}{\sqrt\beta}\Big)^{J_{k-1}+n_k-1} t^{J_{k-1}-1}\,\tau^{n_k}\,,\end{aligned}$$

hence finally

$$\Big\|R_N^{0,K}\Big\|_{L^\infty(\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\leq C\mu_N\beta^{d/2}\sum_{k=1}^K\exp\Big(k(k+1)+2^k\log\gamma\Big)\,,$$

and the proposition is proved. □

5. Proof of the convergence

Our goal in this section is to prove that in the scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$

$$\|\lambda^d f_N^{(1)}-g_N\|_{L^\infty(\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\to 0\,,\quad \text{as } N\to\infty\,.$$

Thanks to Proposition 4.2 we are reduced to proving that, with the notation introduced in (4.18) and (4.20),

$$\|\lambda^d f_N^{(1,K)}-g_N^{(1,K)}\|_{L^\infty(\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\to 0\,,\quad \text{as } N\to\infty\,.$$

5.1. Energy and time truncations.

5.1.1. *Energy truncation.* We introduce a (large) parameter $E>0$ and we define

$$(5.1)\qquad \begin{aligned} f_{N,E}^{(1,K)} &:= \sum_{j_1=1}^{n_1-1}\cdots\sum_{j_K=0}^{n_K-1} Q_{1,J_1}(\tau)Q_{J_1,J_2}(\tau)\dots Q_{J_{K-1},J_K}(\tau)\mathbf{1}_{H_{J_K}(Z_{J_K})\leq E^2}\, f_N^{0(J_K)}\,,\\ g_{N,E}^{(1,K)} &:= \sum_{j_1=1}^{n_1-1}\cdots\sum_{j_K=0}^{n_K-1} Q^0_{1,J_1}(\tau)Q^0_{J_1,J_2}(\tau)\dots Q^0_{J_{K-1},J_K}(\tau)\mathbf{1}_{H_{J_K}(Z_{J_K})\leq E^2}\, g_N^{0(J_K)}\,.\end{aligned}$$

Then, we have the following error estimate.

Proposition 5.1. *There is a constant C depending only on β and d such that, as N goes to infinity, in the scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$, the following bounds hold:*

$$\begin{aligned}\|\lambda^d(f_N^{(1,K)}-f_{N,E}^{(1,K)})\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)}+\|g_N^{(1,K)}-g_{N,E}^{(1,K)}\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\\ \leq 2^{K(K+1)}(Ct)^{2^{K+1}}\mu_N^2 e^{-\frac\beta4 E^2}\,,\end{aligned}$$

with $K=t^2/c\gamma$.

Proof. We have obviously

$$(5.2)\qquad \|\mathbf{1}_{H_{J_K}(Z_{J_K})\geq E^2}\, f_N^{(J_K)}(0)\|_{\varepsilon,J_K,\beta/2}\leq \|f_N^{(J_K)}(0)\|_{\varepsilon,J_K,\beta}\, e^{-\frac\beta4 E^2}\,.$$

On the other hand, by Lemma 4.1, we get

$$\begin{aligned}&\lambda^d\Big\|Q_{1,J_1}(\tau)Q_{J_1,J_2}(\tau)\dots Q_{J_{K-1},J_K}(\tau)\mathbf{1}_{H_{J_K}(Z_{J_K})\geq E^2}\,f_N^{(J_K)}(0)\Big\|_{L^\infty(\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\\&\leq\lambda^d\Big\||Q|_{1,J_K}(t)\mathbf{1}_{H_{J_K}(Z_{J_K})\geq E^2}\,f_N^{(J_K)}(0)\Big\|_{L^\infty(\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\\&\leq\Big(\frac{C_d\lambda^d t}{(\beta/2)^{(d+1)/2}}\Big)^{J_K-1}\lambda^d\|\mathbf{1}_{H_{J_K}(Z_{J_K})\geq E^2}f_N^{(J_K)}(0)\|_{\varepsilon,J_K,\beta/2}\,.\end{aligned}$$

It follows that

$$\begin{aligned}&\lambda^d\Big\|Q_{1,J_1}(\tau)Q_{J_1,J_2}(\tau)\dots Q_{J_{K-1},J_K}(\tau)\mathbf{1}_{H_{J_K}(Z_{J_K})\geq E^2}\,f_N^{(J_K)}(0)\Big\|_{L^\infty(\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\\&\leq\mu_N^2(Ct)^{2^{K+1}}e^{-\frac{\beta}{4}E^2}\end{aligned}$$

thanks to (5.2), (4.12) and to the fact that $J_K\leq\mathcal{N}_K\leq 2^{K+1}$. A similar estimate holds in the Boltzmann case. Summing over all possible choices of j_k proves the lemma. □

5.1.2. *Time separation.* We choose a small parameter $\delta>0$, such that $n_K\delta\ll\tau$ and further restrict the study to the case when $t_i-t_{i+1}\geq\delta$. Given n, we first introduce

$$\mathcal{T}_n(\tau):=\Big\{T_n=(t_1,\dots,t_n)\,/\,t_i<t_{i-1}\ \text{ with }\ t_{n+1}=0\,,\,t_0=\tau\Big\}\,.$$

The time separation is taken into account by the subset

$$\mathcal{T}_{n,\delta}(\tau):=\Big\{T_n\in\mathcal{T}_n(\tau)\,/\,t_i-t_{i+1}\geq\delta\Big\}\,.\tag{5.3}$$

Recall that $J_k=j_1+\dots+j_k$ and $J_0=1$. For $1\leq k\leq K$, we then define

$$Q^\delta_{J_{k-1},J_k}(\tau):=\int_{T_{j_k}\in\mathcal{T}_{j_k,\delta}(\tau)}\mathbf{S}_{J_{k-1}}(\tau-t_1)C_{J_{k-1},J_{k-1}+1}\dots\mathbf{S}_{J_k}(t_{j_k})\,dt_{j_k}\dots dt_1$$

and

$$Q^{0,\delta}_{J_{k-1},J_k}(\tau):=\int_{T_{j_k}\in\mathcal{T}_{j_k,\delta}(\tau)}\mathbf{S}^0_{J_{k-1}}(\tau-t_1)C^0_{J_{k-1},J_{k-1}+1}\dots\mathbf{S}^0_{J_k}(t_{j_k})\,dt_{j_k}\dots dt_1\,.$$

We set accordingly

$$\begin{aligned}f^{(1,K)}_{N,E,\delta}&:=\sum_{j_1=1}^{n_1-1}\dots\sum_{j_K=0}^{n_K-1}Q^\delta_{1,J_1}(\tau)Q^\delta_{J_1,J_2}(\tau)\dots Q^\delta_{J_{K-1},J_K}(\tau)\mathbf{1}_{H_{J_K}(Z_{J_K})\leq E^2}\,f_N^{0(J_K)}\,,\\g^{(1,K)}_{N,E,\delta}&:=\sum_{j_1=1}^{n_1-1}\dots\sum_{j_K=0}^{n_K-1}Q^{0,\delta}_{1,J_1}(\tau)Q^{0,\delta}_{J_1,J_2}(\tau)\dots Q^{0,\delta}_{J_{K-1},J_K}(\tau)\mathbf{1}_{H_{J_K}(Z_{J_K})\leq E^2}\,g_N^{0(J_K)}\,.\end{aligned}\tag{5.4}$$

Then the following holds.

Proposition 5.2. *There is a constant C depending only on β and d such that, as N goes to infinity, in the scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv1$, the following bounds hold:*

$$\begin{aligned}&\|\lambda^d(f^{(1,K)}_{N,E}-f^{(1,K)}_{N,E,\delta})\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)}+\|g^{(1,K)}_{N,E}-g^{(1,K)}_{N,E,\delta}\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\\&\qquad\leq2^{K(K+1)}(Ct)^{2^{K+1}}\mu_N^2\frac{2^{2K}\delta}{t}\,,\end{aligned}$$

with $K=t^2/c\gamma$ (with the notation of Proposition 4.2).

Proof. The proof follows exactly the same lines as the previous one. The point is to see that the following continuity estimate holds:

$$\Big\| \big(Q_{1,n}(t) - Q^\delta_{1,n}(t)\big) f_n \Big\|_{L^\infty([0,t]\times \mathbf{T}^d_\lambda \times \mathbf{R}^d)} \leq \left(\frac{C\lambda^d t}{\beta^{\frac d2 +1}} \right)^{n-1} n^2 \frac{\delta}{t} \|f_n\|_{\varepsilon,n,\beta}$$

since the integration with respect to time provides $\delta t^{n-1}/(n-1)!$ instead of $t^n/n!$, and there are n possible choices of the integral to be modified. Of course a similar estimate holds in the Boltzmann case. Summing over all possible choices of j_k proves therefore the lemma. □

5.1.3. *Reformulation in terms of pseudo-trajectories.* Each elementary term of the sums defining $f^{(1,K)}_{N,E,\delta}$ and $g^{(1,K)}_{N,E,\delta}$ given by (5.4) can be associated with some "pseudo-trajectories". As explained in [22, 9, 14], in this formulation the characteristics associated with the operators $\mathbf{S}_{s+i}(t_i - t_{i+1})$ and $\mathbf{S}^0_{s+i}(t_i - t_{i+1})$ are followed backwards in time between two consecutive times t_{i+1} and t_i, and collision terms (associated with $C_{s+i,s+i+1}$ and $C^0_{s+i,s+i+1}$) are seen as source terms in which "additional particles" are "adjoined" to the system.
The main heuristic idea is then that for the BBGKY hierarchy, in the time interval $[t_{i+1}, t_i]$ between two collisions $C_{s+i-1,s+i}$ and $C_{s+i,s+i+1}$, the particles should not interact in general so trajectories should correspond to the free flow $\mathbf{S}^0_{s+i}$. In order to prevent recollisions in the time interval $[t_{i+1}, t_i]$, some bad sets in phase space must be removed. A geometrical control of free trajectories in the torus (stated in Lemma 5.3) makes it possible, exactly as in [14], to define those small bad sets, outside of which the flow between two collision times is indeed free flow (see Proposition 5.4).

5.2. **Reduction to non pathological trajectories.**

We shall fix two parameters $\bar a, \varepsilon_0$ such that

$$2^K \varepsilon \ll \bar a \ll \varepsilon_0 \ll \min(\delta E, \lambda)\,. \tag{5.5}$$

We recall that the parameter δ (introduced in Paragraph 5.1.2) scales like time while E (introduced in Paragraph 5.1.1), scales like a velocity. The parameters $\bar a$ and ε_0, just like ε, will have the scaling of a distance. We define $B_r(x)$ the ball centered at x of radius r.

The basic point to prove that recollisions have asymptotically negligible probability is the following control on free trajectories. We denote by d the distance on the torus.

Lemma 5.3. *Given $t > 0$, $\lambda \geq 1$ and $\bar a > 0$ satisfying $2^K\varepsilon \ll \bar a \ll \varepsilon_0 \ll \lambda$, consider $\bar x_1, \bar x_2$ in $\mathbf{T}^d_\lambda$ such that $d(\bar x_1, \bar x_2) \geq \varepsilon_0$, and $v_1 \in \mathbf{R}^d$ such that $|v_1| \leq E$. Then there exists a subset $K(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)$ of $\mathbf{R}^d$ defined as the intersection of a sphere and a finite union of cones, the measure of which satisfies*

$$|K(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)| \leq CE^d \left(\Big(\frac{\bar a}{\varepsilon_0}\Big)^{d-1} + \Big(E\frac{t}{\lambda}\Big)^d \Big(\frac{\bar a}{\lambda}\Big)^{d-1} \right)$$

and a subset $K_\delta(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)$ of $\mathbf{R}^d$, the measure of which satisfies

$$|K_\delta(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)| \leq CE \left(\Big(\frac{\varepsilon_0}{\delta}\Big)^{d-1} + \Big(E\frac{t}{\lambda}\Big)^d E^{d-1} \Big(\frac{\bar\varepsilon_0}{\lambda}\Big)^{d-1} \right)$$

such that for any $x_1 \in B_{\bar a}(\bar x_1)$, any $x_2 \in B_{\bar a}(\bar x_2)$ and any $v_2 \in B_E(0)$, the following results hold.

- *If $v_1 - v_2 \notin K(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)$, then*

$$\forall s \in [0,t]\,, \quad d(x_1 - v_1 s, x_2 - v_2 s) > \varepsilon\,;$$

- *If $v_1 - v_2 \notin K_\delta(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)$*

$$\forall s \in [\delta,t]\,, \quad d(x_1 - v_1 s, x_2 - v_2 s) > \varepsilon_0\,.$$

The proof of this lemma is a simple adaptation of Lemma 12.2.1 in [14], and is given in Appendix B. Note that this is the only point of the convergence proof which differs in the case of the torus $\mathbf{T}^d_\lambda$ from the case of the whole space $\mathbf{R}^d$. In the case of the torus, there are indeed no longer dispersion properties so waiting for a sufficiently long time, we expect trajectories to go back ε-close to their initial positions. This return time can be even smaller for a small set of periodic or quasi-periodic trajectories.

5.2.1. *Stability of good configurations.* The proof of convergence then follows the arguments of [14]. The strategy is to slightly modify (in a uniform way) the collision integrals defining $f^{(1,K)}_{N,E,\delta}$ and $g^{(1,K)}_{N,E,\delta}$ such that the remaining pseudo-trajectories can be decomposed as a succession of free transport and binary collisions, without any recollision.
More precisely, we will define recursively some bad sets of velocities and deflection angles to be truncated from the integration domain. The bad set $\mathcal{B}_k$ to be eliminated in the k-th collision integral is defined in terms of the configuration of k particles at time t_k (which will be a good configuration by construction), and therefore depends on the sequence of collision times $(t_1, t_2, \dots t_k)$, as well as on the initial state z_1 and of the velocities, deflection angles and parameters $(m_i, v_{i+1}, \nu_{i+1})_{i<k}$ describing the previous collisions.
Note that initially the system consists of only one particle and therefore is a good configuration.

At this stage, we only describe one elementary step of the construction. The following proposition shows indeed how to eliminate bad sets in phase space so that for any particle outside such bad sets, adjoined to a good configuration, the resulting configuration is again a good configuration in the following sense: we define, for any $c > 0$ and $k \in \mathbf{N}^*$,

$$\mathcal{G}_k(c) = \Big\{ Z_k \in \mathbf{T}^{dk}_\lambda \times \mathbf{R}^{dk} \,/\, \forall s \in [0,t], \quad \forall i \neq j, \quad d(x_i - sv_i, x_j - v_j) \geq c \Big\} \,.$$

Proposition 5.4 ([14])**.** *Let $\bar a, \varepsilon_0$ satisfy (5.5). Given $\overline{Z}_k \in \mathcal{G}_k(\varepsilon_0)$, $m \leq k$, there is a subset $\mathcal{B}^m_k(\overline{Z}_k)$ of $\mathbf{S}^{d-1}_1 \times B_E(0)$ of small measure*

$$\big|\mathcal{B}^m_k(\overline{Z}_k)\big| \leq Ck \left(E^d \left(\frac{\bar a}{\varepsilon_0}\right)^{d-1} + E^d \Big(E\frac{t}{\lambda}\Big)^d \Big(\frac{\varepsilon_0}{\lambda}\Big)^{d-1} + E\left(\frac{\varepsilon_0}{\delta}\right)^{d-1} \right) , \tag{5.6}$$

such that good configurations close to $\overline{Z}_k$ are stable by adjunction of a collisional particle close to the particle $\bar x_m$ and not belonging to $\mathcal{B}_k(\overline{Z}_k)$, in the following sense.

Consider $(\nu, v) \in (\mathbf{S}^{d-1}_1 \times B_E(0)) \setminus \mathcal{B}_k(\overline{Z}_k)$ and let Z_k be a configuration of k particles such that $V_k = \overline{V}_k$ and $|X_k - \overline{X}_k| \leq \bar a$. The new particle is added at $x_m + \varepsilon\nu$ and has velocity v.

- *If $\nu \cdot (v - \bar v_m) < 0$ then for all $\varepsilon > 0$ sufficiently small,*

$$\forall s \in]0,t] \,, \quad \begin{cases} \forall i \neq j \in [1,k] \,, \quad d(x_i - s\bar v_i, x_j - s\bar v_j) > \varepsilon \,, \\ \forall j \in [1,k] \,, \quad d(x_m + \varepsilon\nu - sv, x_j - s\bar v_j) > \varepsilon \,. \end{cases} \tag{5.7}$$

Moreover after the time δ, the $k+1$ particles are in a good configuration:

$$\forall s \in [\delta,t] \,, \quad \begin{cases} (X_k - sV_k, V_k, x_m + \varepsilon\nu - sv, v) \in \mathcal{G}_{k+1}(\varepsilon_0/2) \\ (\bar X_k - sV_k, V_k, \bar x_m - sv, v) \in \mathcal{G}_{k+1}(\varepsilon_0) \,. \end{cases} \tag{5.8}$$

- *If* $\nu \cdot (v - \bar v_m) > 0$ *then*

$$
\forall s \in]0,t]\,,\quad \begin{cases} \forall i \neq j \in [1,k] \setminus \{m\}\,, \quad d(x_i - s\bar v_i, x_j - s\bar v_j) > \varepsilon\,,\\ \forall j \in [1,k] \setminus \{m\}\,, \quad d(x_m + \varepsilon\nu - sv^*, x_j - s\bar v_j) > \varepsilon\,,\\ \forall j \in [1,k] \setminus \{m\}\,, \quad d(x_m - sv^*_m, x_j - s\bar v_j) > \varepsilon\,,\\ d(x_m - sv^*_m, x_m + \varepsilon\nu - sv^*) > \varepsilon\,. \end{cases} \tag{5.9}
$$

Moreover after the time δ*, the* $k+1$ *particles are in a good configuration:*

$$
\forall s \in [\delta, t]\,,\quad \begin{cases} (\{x_j - sv_j, v_j\}_{j\neq m}, x_m - sv^*_m, v^*_m, x_m + \varepsilon\nu - sv^*, v^*) \in \mathcal{G}_{k+1}(\varepsilon_0/2)\\ (\{\bar x_j - sv_j, v_j\}_{j\neq m}, \bar x_m - sv^*_m, v^*_m, \bar x_m - sv^*, v^*) \in \mathcal{G}_{k+1}(\varepsilon_0)\,. \end{cases} \tag{5.10}
$$

5.2.2. *Non pathological trajectories.* Now we are ready to define non-pathological trajectories.

Given $z \in \mathbf{T}^d_\lambda \times B_E(0)$, we call $z^0(\tau)$ the position of the backward free flow initiated from z, at time $t_1 \leq \tau \leq t$. Then given a deflection angle ν_2 and a velocity v_2 we call $Z^0_2(\tau)$ the position at time $t_2 \leq \tau < t_1$ of the Boltzmann pseudo-trajectory initiated by the adjunction of the particle (ν_2, v_2) to the particle $z^0(t_1)$ (which is simply free-flow in the pre-collisional case $(v_2 - v_1)\cdot\nu_2 < 0$, and free-flow after scattering of particles $z^0(t_1)$ and (ν_2, v_2) in the post-collisional case $(v_2 - v_1)\cdot\nu_2 > 0$).

By induction given $z \in \mathbf{T}^d_\lambda \times B_E(0)$, collision times $T = (t_1, \dots, t_{J_K-1})$ and labels of the collision particles $m = (m_1, \dots, m_{J_K-1})$ as well as a collection of velocities $v_1, \dots, v_{J_K}$ and deflection angles $\nu_1, \dots, \nu_{J_K}$, we denote for each $1 \leq k \leq J_K - 1$ by $Z^0_{k+1}(\tau)$ the position at time $t_{k+1} \leq s < t_k$ of the pseudo-trajectory initiated by the adjunction of the particle (ν_{k+1}, v_{k+1}) to the particle $z^0_{m_k}(t_k)$ (which is simply free-flow in the pre-collisional case $(v_{k+1} - v_{m_{k+1}})\cdot\nu_{k+1} < 0$, and free-flow after scattering of particles $z^0_{m_k}(t_k)$ and (ν_{1+k}, v_{1+k}) in the post-collisional case $(v_{k+1} - v_{m_{k+1}})\cdot\nu_{k+1} > 0$). Notice that $s \mapsto Z^0_{k+1}(s)$ is pointwise right-continuous on $[0, t_k]$.

In the same way, given $z \in \mathbf{T}^d_\lambda \times B_E(0)$, collision times $T = (t_1, \dots, t_{J_K})$ and labels of the collision particles $m = (m_1, \dots, m_{J_K})$, as well as a collection of velocities $v_1, \dots, v_{J_K}$ and deflection angles $\nu_1, \dots, \nu_{J_K}$, we define the BBGKY pseudo-trajectories by

- the backward dynamics of the system of k particles on the interval $]t_k, t_{k-1}]$;
- the adjunction of the $(k+1)$-th particle at time t_k with position $x_{m_k}(t_k) + \varepsilon\nu_{k+1}$ and velocity v_{k+1}.

Note that, at this stage, this pseudo-trajectory differs from the Boltzmann pseudo-trajectory on the one hand because of micro-translations at each collision time t_k, and on the other hand because of possible recollisions on intervals $]t_k, t_{k-1}[$.

Using the notation introduced above, we shall write $g^{(1,K)}_{N,E,\delta}$ in terms of trajectories issued from $z \in \mathbf{T}^d_\lambda \times B_E(0)$. Given $J = (j_1, \dots, j_K)$, we introduce a decreasing sequence of J_K times with j_k terms in the k^{th} time interval

$$
\mathcal{T}_{J,\delta}(\tau) := \Big\{ T_{J_k} = (t_1, \dots, t_{J_K-1}) \,/\, t_i < t_{i-1} - \delta; \quad (t_{J_k}, \dots, t_{J_{k-1}+1}) \in [t - k\tau, t - (k-1)\tau] \Big\}
$$

We can write (setting $j_0 := 0$)

$$
g^{(1,K)}_{N,E,\delta}(t,x,v) = \sum_{j_1=1}^{n_1-1} \cdots \sum_{j_K=0}^{n_K-1} \sum_{m \in \{1,\dots,J_K\}} G^{(1,K)}_{N,E,\delta}(J,m)(t,x,v)
$$

with

$$G^{(1,K)}_{N,E,\delta}(J,m)(t,x,v) := \int_{\mathcal{T}_{J,\delta}(t)} dT_{J_K} \int_{\mathbf{S}_1^{d-1}\times B_E(0)} d\nu_2 dv_2((v_2 - v^0_{m_1}(t_1))\cdot\nu_2)$$
$$\dots \int_{\mathbf{S}_1^{d-1}\times B_E(0)} d\nu_{J_K} dv_{J_K}((v_{J_K} - v^0_{m_{J_K-1}}(t_{J_K-1}))\cdot\nu_{J_K}) \mathbf{1}_{H_{J_K}(Z^0_{J_K}(0))\le E^2} g_N^{0(J_K)}(Z^0_{J_K}(0))\,.$$

Thanks to Proposition 5.4, we may define recursively the approximate Boltzmann first marginal

$$\widetilde{g}^{(1,K)}_{N,E,\delta}(t,x,v) := \sum_{j_1=1}^{n_1-1}\cdots\sum_{j_K=0}^{n_K-1}\sum_{m\in\{1,\dots,J_K\}} \widetilde{G}^{(1,K)}_{N,E,\delta}(J,m)(t,x,v)\,,$$

where

$$\begin{aligned}\widetilde{G}^{(1,K)}_{N,E,\delta}(J,m)(t,x,v) &:= \int_{\mathcal{T}_{J,\delta}(t)} dT_{J_K} \int_{{}^c\mathcal{B}_1^{m_1}(z^0(t_1))} d\nu_2 dv_2((v_2 - v^0_{m_1}(t_1))\cdot\nu_2)\\ &\dots \int_{{}^c\mathcal{B}^{m_{J_K-1}}_{J_K-1}(Z^0_{J_K-1}(t_{J_K-1}))} d\nu_{J_K} dv_{J_K}((v_{J_K} - v^0_{m_{J_K-1}}(t_{J_K-1}))\cdot\nu_{J_K})\\ &\qquad\times\mathbf{1}_{H(Z^0_{J_K}(0))\le E^2}\; g_N^{0(J_K)}(Z^0_{J_K}(0))\,.\end{aligned} \tag{5.11}$$

Similarly, we define the approximate BBGKY first marginal, by *excluding the same bad set of velocities* $\mathcal{B}_1^{m_1}(z^0(t_1)),\dots,\mathcal{B}^{m_{J_K-1}}_{J_K-1}(Z^0_{J_K-1}(t_{J_K-1}))$ as in the Boltzmann case at each step: taking into account that the prefactors in the collision operators $(N-k)\varepsilon^{d-1}$ are of the order λ^d, we define the approximate first marginal

$$\lambda^d \widetilde{f}^{(1,K)}_{N,E,\delta}(t,x,v) = \sum_{j_1=1}^{n_1-1}\cdots\sum_{j_K=0}^{n_K-1}\sum_{m\in\{1,\dots,J_K\}} \widetilde{F}^{(1,K)}_{N,E,\delta}(J,m)(t,x,v)\,,$$

where

$$\begin{aligned}\widetilde{F}^{(1,K)}_{N,E,\delta}(J,m)(t,x,v) &:= \frac{(N-1)!}{(N-J_K)!}(\varepsilon^{d-1}\lambda^{-d})^{J_K-1}\int_{\mathcal{T}_{J,\delta}(t)} dT_{J_K} \int_{{}^c\mathcal{B}_1^{m_1}(z^0(t_1))} d\nu_2 dv_2((v_2 - v_{m_1}(t_1))\cdot\nu_2)\\ &\dots \int_{{}^c\mathcal{B}^{m_{J_K-1}}_{J_K-1}(Z^0_{J_K-1}(t_{J_K-1}))} d\nu_{J_K} dv_{J_K}((v_{J_K} - v_{m_{J_K-1}}(t_{J_K-1}))\cdot\nu_{J_K})\\ &\qquad\times\mathbf{1}_{H(Z_{J_K}(0))\le E^2}\; \lambda^{dJ_K} f_N^{0(J_K)}(Z_{J_K}(0))\,.\end{aligned} \tag{5.12}$$

Note that (5.12) differs from (5.11) by

- the initial configuration $Z_{J_K}(0)$ instead of $Z^0_{J_K}(0)$;
- the initial data $\lambda^{dJ_K} f_N^{0(J_K)}$ instead of $g_N^{0(J_K)}$;
- and the prefactor

$$\frac{(N-1)!}{(N-J_K)!}(\varepsilon^{d-1}\lambda^{-d})^{J_K-1}\,.$$

As a consequence of Proposition 5.4 and of the continuity estimates in Lemma 4.1, we have the following result.

Proposition 5.5. *Let $\bar a, \varepsilon_0 \ll 1$ satisfying* (5.5)*. There is a constant C depending only on β and d such that, as N goes to infinity, in the scaling $N\varepsilon^{d-1}\lambda^{-d} \equiv 1$, the following holds:*

$$\Big\| g^{(1,K)}_{N,E,\delta} - \widetilde g^{(1,K)}_{N,E,\delta}\Big\|_{L^\infty([0,t]\times \mathbf{T}^d_\lambda\times\mathbf{R}^d)} + \Big\| \lambda^d\big(f^{(1,K)}_{N,E,\delta} - \widetilde f^{(1,K)}_{N,E,\delta}\big)\Big\|_{L^\infty([0,t]\times \mathbf{T}^d_\lambda\times\mathbf{R}^d)}$$
$$\leq C\mu_N^2 2^{K(K+1)} (Ct)^{2^{K+1}} 2^{2K} \left(E^d \left(\frac{\bar a}{\varepsilon_0}\right)^{d-1} + E^d\Big(E\frac{t}{\lambda}\Big)^d \Big(\frac{\varepsilon_0}{\lambda}\Big)^{d-1} + E\left(\frac{\varepsilon_0}{\delta}\right)^{d-1}\right).$$

Proof. The proof follows exactly the same lines as the proof of Propositions 5.1 and 5.2. In the usual continuity estimate for the elementary collision operator, the integration with respect to velocity brings a factor $(2\pi/\beta)^{d/2}$, while removing the integration over the pathological set $\mathcal{B}_k^{m_k}$ gives an error

$$Ck \left(E^d \left(\frac{\bar a}{\varepsilon_0}\right)^{d-1} + E^d\Big(E\frac{t}{\lambda}\Big)^d \Big(\frac{\varepsilon_0}{\lambda}\Big)^{d-1} + E\left(\frac{\varepsilon_0}{\delta}\right)^{d-1}\right)$$

according to Proposition 5.4.
For each given J, there are $J_K \leq 2^K$ possible choices of the integral to be modified. Therefore,

$$\Big\| g^{(1,K)}_{N,E,\delta} - \widetilde g^{(1,K)}_{N,E,\delta}\Big\|_{L^\infty([0,t]\times \mathbf{T}^d_\lambda\times\mathbf{R}^d)} + \Big\| \lambda^d\big(f^{(1,K)}_{N,E,\delta} - \widetilde f^{(1,K)}_{N,E,\delta}\big)\Big\|_{L^\infty([0,t]\times \mathbf{T}^d_\lambda\times\mathbf{R}^d)}$$
$$\leq C\mu_N^2 (Ct)^{2^{K+1}} 2^{2K} \left(E^d \left(\frac{\bar a}{\varepsilon_0}\right)^{d-1} + E^d\Big(E\frac{t}{\lambda}\Big)^d \Big(\frac{\varepsilon_0}{\lambda}\Big)^{d-1} + E\left(\frac{\varepsilon_0}{\delta}\right)^{d-1}\right)$$

where as previously C depends only on d and β.
Finally we have to sum over all the possible choices of J, which provides the additional factor $2^{K(K+1)}$ in the estimate. □

5.3. **Proximity of non pathological trajectories.** This paragraph is devoted to the proof, by induction, that the BBGKY and Boltzmann pseudo-trajectories remain close for all times, in particular that there is no recollision for the BBGKY dynamics.

The following statement is analogous to Lemma 14.1.1 of [14]. We recall that the notation $Z^0(t)$ and $Z(t)$ were defined in Paragraph 5.2.2.

Lemma 5.6. *Fix $J = (j_1, \dots, j_k)$, $T \in \mathcal{T}_{J,\delta}(t)$. Consider for all $i \in \{1, \dots J_K\}$, a label m_{i+1}, an impact parameter ν_{i+1} and a velocity v_{i+1} such that $(\nu_{i+1}, v_{i+1}) \notin \mathcal{B}_i^{m_i}(Z_i^0(t_i))$. Then, for ε sufficiently small, for all $i \in [1, J_K]$,*

$$\forall \ell \leq i+1, \qquad |x_\ell(t_{i+1}) - x^0_\ell(t_{i+1})| \leq \varepsilon i \quad \text{and} \quad v_\ell(t_{i+1}) = v^0_\ell(t_{i+1})\,. \tag{5.13}$$

Proof. We proceed by induction on i, the index of the time variables t_{i+1} for $0 \leq i \leq J_K$.

We first notice that by construction, $z(t_1) - z^0(t_1) = 0$, so (5.13) holds for $i = 0$. The initial configuration containing only one particle, there is no possible recollision!

Now let $k \in [1, J_K]$ be fixed, and assume that (5.13) holds for all $i \leq k$. Let us prove that (5.13) holds for $i = k+1$. We shall consider two cases depending on whether the particle adjoined at time t_i is pre-collisional or post-collisional.

• Let us start with the case of pre-collisional velocities $(v_{i+1}, v_{i+1}(t_i))$ at time t_i. We have for all $s \in [t_{i+1}, t_i]$

$$\begin{aligned} \forall \ell \leq i\,, \quad x^0_\ell(s) &= x^0_\ell(t_i) + (s - t_i) v^0_\ell(t_i)\,, \qquad & v^0_\ell(s) &= v^0_\ell(t_i)\,,\\ x^0_{i+1}(s) &= x^0_{m_i}(t_i) + (s-t_i) v_{i+1}\,, \qquad & v^0_{i+1}(s) &= v_{i+1}\,. \end{aligned}$$

Now let us study the BBGKY trajectory. We recall that the particle is adjoined in such a way that (ν_{i+1}, v_{i+1}) belongs to ${}^c\mathcal{B}_i^{m_i}(Z_i^0(t_i))$. Provided that ε is sufficiently small, by the induction assumption (5.13) and (5.5), we have

$$\forall \ell \leq i, \quad |x_\ell(t_i) - x_\ell^0(t_i)| \leq \varepsilon(i-1) \leq \bar{a}\,.$$

Since $Z_i^0(t_i)$ belongs to $\mathcal{G}_i(\varepsilon_0)$, we can apply Proposition 5.4 which implies that backwards in time, there is free flow for Z_{i+1}. In particular,

$$\begin{aligned}&\forall \ell < i+1\,, \quad x_\ell(s) = x_\ell(t_i) + (s-t_i)v_\ell(t_i)\,, \qquad v_\ell(s) = v_\ell(t_i)\,,\\ &x_{i+1}(s) = x_{m_i}(t_i) + \varepsilon\nu_{i+1} + (s-t_i)v_{i+1}\,, \qquad v_{i+1}(s) = v_{i+1}\,.\end{aligned} \tag{5.14}$$

We therefore obtain

$$\forall \ell \leq i+1\,, \quad \forall s \in [t_{i+1}, t_i]\,, \quad v_\ell(s) - v_\ell^0(s) = v_\ell(t_i) - v_\ell^0(t_i) = 0\,,$$

and by the induction assumption (5.13) and the shift by ε in (5.14), one has

$$\forall \ell \leq i+1\,, \quad \forall s \in [t_{i+1}, t_i]\,, \quad |x_\ell(s) - x_\ell^0(s)| \leq \varepsilon(i-1) + \varepsilon \leq \varepsilon i\,.$$

• The case of post-collisional velocities $(v_{i+1}, v_{m_i}(t_i))$ at time t_i is very similar. We indeed have $\forall s \in [t_{i+1}, t_i[$

$$\begin{aligned}&\forall \ell < i+1, \quad \ell \neq m_i\,, \quad x_\ell^0(s) = x_\ell^0(t_i) + (s-t_i)v_\ell^{0*}(t_i)\,, \qquad v_\ell^0(s) = v_\ell^0(t_i)\,,\\ &x_{m_i}^0(s) = x_{m_i}^0(t_i) + (s-t_i)v_{m_i}^{0*}(t_i)\,, \qquad v_{m_i}^0(s) = v_{m_i}^{0*}(t_i)\,,\\ &x_{i+1}^0(s) = x_{m_i}^0(t_i) + (s-t_i)v_{i+1}^*\,, \qquad v_{i+1}^0(s) = v_{i+1}^*\,.\end{aligned}$$

Now let us study the BBGKY trajectory. We recall that the particle is adjoined in such a way that (ν_{i+1}, v_{i+1}) belongs to ${}^c\mathcal{B}_i^{m_i}(Z_i^0(t_i))$. Provided that ε is sufficiently small, by the induction assumption (5.13), we have

$$\forall \ell \leq i, \quad |x_\ell(t_i) - x_\ell^0(t_i)| \leq \varepsilon(i-1)\,.$$

Then as in (5.14), we can apply Proposition 5.4 which implies that

$$\begin{aligned}&\forall \ell \leq i+1\,, \quad \forall s \in [t_{i+1}, t_i[\,, \quad v_\ell(s) - v_\ell^0(s) = v_\ell(t_i^-) - v_\ell^0(t_i^-) = 0\,,\\ &\forall \ell \leq i+1\,, \quad \forall s \in [t_{i+1}, t_i[\,, \quad |x_\ell(s) - x_\ell^0(s)| \leq \varepsilon(i-1) + \varepsilon \leq i\varepsilon\,,\end{aligned}$$

which concludes the proof of Lemma 5.6. □

Note that, by construction,

$$Z_{J_K}(0) \in \mathcal{G}_{J_K}(\varepsilon_0/2)\,.$$

In particular, thanks to Lemma 5.6, the formulation (5.12) becomes

$$\begin{aligned}\widetilde{F}_{N,E,\delta}^{(1,K)}(J,m)(t,x,v) = \frac{(N-1)!}{(N-J_K)!}(\varepsilon^{d-1}\lambda^{-d})^{J_K-1}\int_{\mathcal{T}_{J,\delta}(t)} dT_{J_K}\int_{{}^c\mathcal{B}_1^{m_1}(z^0(t_1))} d\nu_2 dv_2((v_2 - v_{m_1}(t_1))\cdot\nu_2)\\ \dots\int_{{}^c\mathcal{B}_{J_K-1}^{m_{J_K-1}}(Z^0_{J_K-1}(t_{J_K-1}))} d\nu_{J_K} dv_{J_K}((v_{J_K} - v_{m_{J_K-1}}(t_{J_K-1}))\cdot\nu_{J_K})\\ \times\mathbf{1}_{H(Z_{J_K}(0))\leq E^2}\mathbf{1}_{Z_{J_K}(0)\in\mathcal{G}_{J_K}(\varepsilon_0/2)}\,\lambda^{dJ_K} f_N^{0(J_K)}(Z_{J_K}(0))\,.\end{aligned} \tag{5.15}$$

As a consequence of the convergence of pseudo-trajectories, we can prove that the approximate Boltzmann first marginal is close to the following distribution:

$$\widetilde{g}_{N,E,\delta,\varepsilon}^{(1,K)}(t,x,v) := \sum_{j_1=1}^{n_1-1}\cdots\sum_{j_K=0}^{n_K-1}\sum_{m\in\{1,\dots,J_K\}}\widetilde{G}_{N,E,\delta,\varepsilon}^{(1,K)}(J,m)(t,x,v)\,,$$

where the initial data in $\tilde G^{(1,K)}_{N,E,\delta,\varepsilon}$ has been artificially changed from $Z^0_{J_K}(0)$ to $Z_{J_K}(0)$

$$
\begin{aligned}
\widetilde G^{(1,K)}_{N,E,\delta,\varepsilon}(J,m)(t,x,v) := \int_{\mathcal{T}_{J,\delta}(t)} dT_{J_K} \int_{{}^c\mathcal{B}_1^{m_1}(z^0(t_1))} d\nu_2 dv_2 ((v_2 - v^0_{m_1}(t_1))\cdot \nu_2) \\
\dots \int_{{}^c\mathcal{B}^{m_{J_K-1}}_{J_K-1}(Z^0_{J_K-1}(t_{J_K-1}))} d\nu_{J_K} dv_{J_K} ((v_{J_K} - v^0_{m_{J_K-1}}(t_{J_K-1}))\cdot \nu_{J_K}) \\
\times \mathbf{1}_{H(Z_{J_K}(0))\leq E^2}\, g_N^{0(J_K)}(Z_{J_K}(0))\,.
\end{aligned}
\tag{5.16}
$$

Note that the only differences left then between (5.15) and (5.16) are the initial data and the prefactor.

Proposition 5.7. *There is a constant C depending only on β and d such that, as N goes to infinity, in the scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$, the following holds:*

$$
\Big\| \widetilde g^{(1,K)}_{N,E,\delta} - \widetilde g^{(1,K)}_{N,E,\delta,\varepsilon} \Big\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)} \leq C\Lambda_N \mu_N^3 2^{K(K+1)} (Ct)^{2^{K+1}} 2^K \varepsilon\,.
$$

Proof. The estimate comes as previously from the continuity of the operator $Q^0_{1,J_K}(t)$ together with the following estimate

$$
\big\| g_N^{0(J_K)}(Z_{J_K}(0)) - g_N^{0(J_K)}(Z^0_{J_K}(0))\big\|_{0,J_K,\beta/2} \leq \big\|\nabla_{Z_{J_K}} g_N^{0(J_K)}\big\|_{0,J_K,\beta/2} \big|Z_{J_K}(0) - Z^0_{J_K}(0)\big|\,.
$$

Since φ^0 is assumed to be Lipschitz continuous, we have

$$
\partial_{x_j} g_N^{0(J_K)} = 0\,,\quad \partial_{v_j} g_N^{0(J_K)} = -\beta v_j g_N^{0(J_K)} \text{ for all } j\neq 1\,,
$$

$$
\partial_{x_1} g_N^{0(J_K)}(Z_{J_K}) = \frac{\partial_{x_1}\varphi^0_N}{\varphi^0_N}(z_1) g_N^{0(J_K)}(Z_{J_K})\,,
$$

$$
\partial_{v_1} g_N^{0(J_K)}(Z_{J_K}) = \Big(-\beta v_1 + \frac{\partial_{v_1}\varphi^0_N}{\varphi^0_N}(z_1)\Big) g_N^{0(J_K)}(Z_{J_K})
$$

so

$$
\big\|\nabla_{Z_{J_K}} g_N^{0(J_K)}\big\|_{0,J_K,\beta/2} \leq C\mu_N\Lambda_N \big\| g_0^{(J_K)}\big\|_{0,J_K,\beta}\,.
$$

Thus we get

$$
\Big\| \widetilde g^{(1,K)}_{N,E,\delta} - \widetilde g^{(1,K)}_{N,E,\delta,\varepsilon} \Big\|_{L^\infty} \leq C\Lambda_N \mu_N^3 (Ct)^{2^{K+1}} 2^K \varepsilon\,.
$$

As previously we have to sum over all the possible choices of J, which provides the additional factor $2^{K(K+1)}$ in the estimate. □

5.4. **Estimate of the main term.** It remains then to compare the approximate BBGKY first marginal $\widetilde f^{(1,K)}_{N,E,\delta}$ defined by (5.15) and the distribution $\widetilde g^{(1,K)}_{N,E,\delta,\varepsilon}$ defined by (5.16). The error, due to the initial data and to the prefactor, can be controlled exactly as in [14] and we end up with the following estimate.

Proposition 5.8. *Let $\bar a, \varepsilon_0, \eta \ll 1$ satisfying* (5.5). *There exists constants c, C depending only on β and d such that, as N goes to infinity, in the scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$, the following* *holds:*

$$\begin{aligned}
&\frac{1}{\mu_N^2}\big\|\lambda^d f_N^{(1)} - g_N\big\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)}\\
&\quad\leq\; C\gamma^2 + 2^{K(K+1)}(Ct)^{2^{K+1}}\left(e^{-\frac{\beta}{4}E^2} + \frac{2^K\delta}{t}\right)\\
&\qquad + C2^{K(K+1)}(Ct)^{2^{K+1}}2^{2K}\left(E^d\left(\frac{\bar a}{\varepsilon_0}\right)^{d-1} + E^d\Big(E\frac{t}{\lambda}\Big)^d\Big(\frac{\varepsilon_0}{\lambda}\Big)^{d-1} + E\left(\frac{\varepsilon_0}{\delta}\right)^{d-1}\right)\\
&\qquad + C2^{K(K+1)}(Ct)^{2^{K+1}}2^{2K}\big(\Lambda_N\mu_N\varepsilon + \varepsilon^{d-1}\lambda^{-d}\big)
\end{aligned}$$

with $K = t^2/c\gamma$, *with the notation of Proposition* 4.2.

Proof. This estimate is obtained essentially by gathering together the estimates in Propositions 4.2, 5.1, 5.2, 5.5 and 5.7. As mentioned previously, there remains to estimate the difference

$$\lambda^d\widetilde f^{(1,K)}_{N,E,\delta} - \widetilde g^{(1,K)}_{N,E,\delta,\varepsilon}\,.$$

Error coming from the initial data: According to Proposition 3.1, we have

$$\Big\|\mathbf{1}_{\mathcal{G}_{J_K}(\varepsilon_0/2)}(\lambda^{dJ_K}f_N^{0(J_K)} - g_N^{0(J_K)})\Big\|_{0,J_K,\beta} \leq C\mu_N^2 J_K\varepsilon\,.$$

Using the continuity estimate in Lemma 4.1, we then deduce that the error due to the initial data can be controlled by

$$C\mu_N^2 2^{K(K+1)}(Ct)^{2^{K+1}}2^K\varepsilon\,.$$

Error coming from the prefactors in the collision operators: In the formula defining the approximate BBGKY first marginal, the elementary collision operators have prefactors of the type

$$(N-k)\varepsilon^{d-1}$$

that we have replaced in the limit by λ^d. For fixed J_K, since

$$\Big(1 - \frac{(N-1)\dots(N-J_K-1)}{N^{J_K-1}}\Big) \leq C\frac{J_K^2}{N}$$

the corresponding error is controlled by

$$C2^{K(K+1)}(Ct)^{2^{K+1}}2^{2K}\frac{1}{N}\,\cdot$$

We therefore conclude that

$$\Big\|\lambda^d\widetilde f^{(1,K)}_{N,E,\delta} - \widetilde g^{(1,K)}_{N,E,\delta}\Big\|_\infty \leq C\mu_N^2 2^{K(K+1)}(Ct)^{2^{K+1}}\Big(2^{2K}\mu_N\Lambda_N\varepsilon + 2^{2K}\varepsilon^{d-1}\lambda^{-d}\Big)\,,$$

which concludes the proof. □

Equipped with all these estimates, we can now optimize the different parameters and prove Theorem 2.2.

Proof of Theorem 2.2. We choose for the parameters the following orders of magnitude:

$$\delta\sim\varepsilon^{\frac{d-1}{d+1}}\,,\quad \varepsilon_0\sim\varepsilon^{\frac{d}{d+1}}\,,\quad E\sim\sqrt{|\log\varepsilon|}\,,\quad K\sim\log|\log\varepsilon|\,.$$

Then Proposition 5.8 implies (since $t\geq 1$ and $\lambda\geq 1$) that

$$\begin{aligned}
\big\|\lambda^d f_N^{(1)} - g_N\big\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)} &\leq C\mu_N^2\left(\gamma^2 + (Ct)^{2^{K+1}}\big(\varepsilon^{\frac{d-1}{d+1}}|\log\varepsilon|^d + \varepsilon^{d-1}\lambda^{-d}\big)\right)\\
&\leq C\mu_N^2\left(\frac{t^2}{\log\log N}\right)^2\,,
\end{aligned}$$

where we used that $\gamma = t^2/CK$ (see (4.23)) and $t = o(\sqrt{\log\log N})$. □

6. Proof of the diffusive limit : proof of Theorem 2.3

In Theorem 2.2, we have shown that the tagged particle distribution $\lambda_N^d f_N^{(1)}$ remains close to the solution $g_N(t,x,v)$ of the linear-Boltzmann equation on $\mathbf{T}^d_{\lambda_N} \times \mathbf{R}^d$

$$\begin{aligned} &\partial_t g_N + v\cdot\nabla_x g_N = -Lg_N ,\\ &Lg_N(v) := \iint_{\mathbf{S}^{d-1}\times\mathbf{R}^d} [g_N(v)M_\beta(v_1) - g_N(v')M_\beta(v_1')] \big((v-v_1)\cdot\nu\big)_+ \, d\nu dv_1 \,. \end{aligned} \tag{6.1}$$

More generally, our proof implies that the whole trajectory of the tagged particle $\{x_1(s)\}_{s\leq T\lambda_N^2}$ can be approximated with high probability by the trajectory of $\{x^0_1(s)\}_{s\leq T\lambda_N^2}$ (see Lemma 5.6). The latter process is much simpler to study as its velocities are given by a Markov process. These two points of view lead to two strategies to derive Theorem 2.3. We will first present an analytic approach to show that g_N can be approximated at large scale by the diffusion (2.16). Then we turn to an alternative method to show the convergence of the rescaled trajectory to a brownian which will rely on probabilist estimates for $\{x^0_1(s)\}_{s\leq T\lambda_N^2}$.

In the following the macroscopic space variable will be denoted by $y \in \mathbf{T}^d$ and the macroscopic time by $\tau \in [0,T]$ (not to be confused with the τ used in the previous sections).

6.1. **Convergence to the heat equation.** We are going to derive the convergence of $\lambda_N^d f_N^{(1)}$ to (2.16). This boils down to proving that g_N can be approximated at large scale by a diffusion which is a standard procedure (see [3]). For the sake of completeness, we recall the salient features of the proof and sketch how they can be adapted to our framework.

Step 1: Approximation by the linear Boltzmann equation.
The scaling (2.15) of the initial data $\rho_{N,0}(x) = \lambda_N^{d\zeta}\ \chi(\lambda_N^{\zeta-1}x)$ leads to

$$\mu_N = \lambda_N^{d\zeta}\,, \qquad \Lambda_N = \lambda_N^{(d+1)\zeta-1} \ll N^{\frac{2}{d^2-1}}\,. \tag{6.2}$$

Using (2.12), we get after a diffusive rescaling that for $T>0$

$$\begin{aligned} \big\|\lambda_N^d f_N^{(1)}(\lambda_N^2\tau,\lambda_N y,v) - g_N(\lambda_N^2\tau,\lambda_N y,v)\big\|_{L^\infty([0,T]\times\mathbf{T}^d\times\mathbf{R}^d)} &\leq C\left(\frac{\lambda_N^{d\zeta+4}}{\log\log N}\right)^2 T^4\\ &\leq C\frac{T^4}{\big(\log\log N\big)^{1/10}}\,, \end{aligned} \tag{6.3}$$

where we used that $\zeta \in (0,\frac{1}{4d})$ and $\lambda_N = o((\log\log N)^{1/5})$. Note that the bounds on the exponents here are not sharp.

Let F be a smooth function in $\mathbf{T}^d$. The expectation with respect to the rescaled tagged particle (2.14) reads

$$\begin{aligned} \mathbb{E}\left(F\left(\Xi(\tau)\right)\right) &= \int_{\mathbf{T}^d_{\lambda_N}\times\mathbf{R}^d} dxdvF\left(\frac{x}{\lambda_N}\right) f_N^{(1)}(\lambda_N^2\tau,x,v)\\ &= \int_{\mathbf{T}^d\times\mathbf{R}^d} dydvF\left(y\right)\ \lambda_N^d f_N^{(1)}(\lambda_N^2\tau,\lambda_N y,v)\\ &= \int_{\mathbf{T}^d} dyF\left(y\right)\ \int_{\mathbf{R}^d} dv\ g_N(\lambda_N^2\tau,\lambda_N y,v) + O\left(\frac{1}{\big(\log\log N\big)^{1/10}}\right)\ , \end{aligned}$$

where we used (6.3) in the last equation. Thus it is enough to focus on the solution of the linear Boltzmann equation and show that the rescaled density

$$\forall \tau, y \in \mathbf{R}^+ \times \mathbf{T}^d, \qquad \rho_N(\tau, y) := \int_{\mathbf{R}^d} g_N\big(\lambda_N^2 \tau, \lambda_N y, v\big)\, dv \tag{6.4}$$

obeys the heat equation in the large N limit. Notice that g_N solves the linear Boltzmann equation with initial data $M_\beta(v)\rho_{N,0}(x)$.

Step 2: Rescaling and diffusion coefficient.
Define the rescaled density

$$\forall \tau, y, v \in \mathbf{R}^+ \times \mathbf{T}^d \times \mathbf{R}^d, \qquad \varphi_N(\tau, y, v) := g_N\big(\lambda_N^2 \tau, \lambda_N y, v\big) M_\beta(v)^{-1}\,, \tag{6.5}$$

which satisfies

$$\partial_\tau \varphi_N + \lambda_N\, v \cdot \nabla_y \varphi_N + \lambda_N^2\, \mathcal{L}\varphi_N = 0\,, \tag{6.6}$$

where $\mathcal{L}$ is the modified operator

$$\mathcal{L}\varphi(v) := \int_{\mathbf{S}^{d-1} \times \mathbf{R}^d} M_\beta(v_1)[\varphi(v) - \varphi(v')]\, ((v - v_1) \cdot \nu)_+ \, d\nu dv_1\,. \tag{6.7}$$

In order to define the diffusion coefficient, we first need to state some properties of the operator $\mathcal{L}$. Let

$$a(v) := \int_{\mathbf{S}^{d-1} \times \mathbf{R}^d} M_\beta(v_1)\, \big((v - v_1) \cdot \nu\big)_+ \, d\nu dv_1\,.$$

Lemma 6.1. *The operator $\mathcal{L}$ is a Fredholm operator of domain $L^2(\mathbf{R}^d, aM_\beta dv)$ and its kernel reduces to the constant functions. In particular, $\mathcal{L}$ is invertible on the set of functions*

$$\Big\{ g \in L^2(\mathbf{R}^d, aM_\beta dv),\ \int g(v)\, M_\beta(v) dv = 0 \Big\}\,.$$

The Lemma follows by noticing the decomposition $\mathcal{L} = a(v)\,\mathrm{Id} - \mathcal{K}$, where Id stands for the identity and $\mathcal{K}$ is a compact operator; see [19] for a proof in the case of the linearized Boltzmann equation, which can be easily adapted to our situation.

We define the vector $b(v) = \big(b_k(v)\big)_{k \le d}$ by

$$\forall k \le d, \qquad \mathcal{L} b_k(v) = v_k\,, \tag{6.8}$$

and the diffusion matrix $D(v) = \big(D_{k,\ell}(v)\big)_{k,\ell \le d}$

$$\forall k, \ell \le d, \qquad \mathcal{L} D_{k,\ell}(v) = v_k b_\ell(v) - \int_{\mathbf{R}^d} v_k b_\ell(v)\, M_\beta(v) dv\,. \tag{6.9}$$

Note that by construction $\int_{\mathbf{R}^d} b_k(v)\, M_\beta(v) dv = \int_{\mathbf{R}^d} D_{k,\ell}\, M_\beta(v) dv = 0$. From the symmetry of the model, one can check (see [10] for instance) that there is a function γ such that

$$b_k(v) = \gamma(|v|) v_k, \quad \text{with} \quad |v|^2 = \sum_{\ell=1}^d v_\ell^2\,.$$

The diffusion coefficient is given by

$$\kappa := \int_{\mathbf{R}^d} v \mathcal{L}^{-1} v\ M_\beta(v) dv = \frac{1}{d} \int_{\mathbf{R}^d} \gamma(|v|) |v|^2\, M_\beta(v) dv\,, \tag{6.10}$$

where we used the symmetry of b to derive the last equality.

Step 3: The diffusive limit.

When λ_N is large, the density (6.4) remains close to the solution of

$$\partial_\tau u_N - \kappa\Delta_y u_N = 0 \text{ on } \mathbf{R}^+ \times \mathbf{T}^d, \qquad u_N(\tau=0,y) = \lambda_N^{d\zeta}\ \chi\big(\lambda_N^\zeta y\big)\,.$$

By construction u_N approximates the solution ρ of the heat equation (2.16).

To prove this convergence, we do not compare directly ρ_N to u_N but instead introduce the Chapman-Enskog expansion (obtained by using the diffusion ansatz and the identification of the powers in $1/\lambda_N$)

$$\Psi_N(\tau,y,v) = u_N(\tau,y) + \frac{1}{\lambda_N} u_{N,1}(\tau,y,v) + \frac{1}{\lambda_N^2} u_{N,2}(\tau,y,v)\,, \tag{6.11}$$

where

$$u_{N,1}(\tau,y,v) := -b(v)\cdot\nabla_y u_N(\tau,y), \quad \text{and} \quad u_{N,2}(\tau,y,v) := \sum_{k,\ell} D_{k,\ell}(v)\ \partial_{y_k}\partial_{y_\ell} u_N(\tau,y)\,.$$

Using (6.8) and (6.9), one can check that Ψ_N is almost a solution of (6.6):

$$\partial_\tau \Psi_N + \lambda_N\ v\cdot\nabla_y\Psi_N + \lambda_N^2\ \mathcal{L}\Psi_N = S_N\,,$$

where the error term S_N is given by

$$S_N = \frac{1}{\lambda_N}\left(\partial_\tau u_{N,1}(\tau,y,v) + \partial_\tau u_{N,2}(\tau,y,v) + \frac{1}{\lambda_N} v\cdot\nabla_y u_{N,2}(\tau,y,v)\right)\,. \tag{6.12}$$

We are going to prove that

$$\lim_{N\to\infty}\int_0^T d\tau \int_{\mathbf{T}^d\times\mathbf{R}^d} dy\ \big(\rho_N(\tau,y) - u_N(\tau,y)\big)^2 = 0\,. \tag{6.13}$$

As $\rho_N(\tau,y) = \int_{\mathbf{R}^d}\ \varphi_N\big(\lambda_N^2\,\tau,\lambda_N y,v\big)\ M_\beta(v)dv$ and applying Jensen inequality, we obtain

$$\int_0^T d\tau \int_{\mathbf{T}^d\times\mathbf{R}^d} dy\ \big(\rho_N(\tau,y) - u_N(\tau,y)\big)^2 \leq \int_0^T d\tau \int_{\mathbf{T}^d\times\mathbf{R}^d} dy M_\beta(v)dv\ \big(\varphi_N(\tau,y,v) - \Psi_N(\tau,y,v)\big)^2\,,$$

where we used notation (6.4) and the fact that for any τ, y

$$\int_{\mathbf{R}^d} dv M_\beta(v) u_{N,1}(\tau,y,v) = \int_{\mathbf{R}^d} dv M_\beta(v) u_{N,2}(\tau,y,v) = 0\,.$$

Thus the proof of the diffusive limit will be complete once we derive the following result.

Lemma 6.2. *For any time $T>0$*

$$\lim_{N\to\infty}\int_0^T d\tau \int_{\mathbf{T}^d\times\mathbf{R}^d} dy M_\beta(v)dv\ (\varphi_N - \Psi_N)^2 = 0\,.$$

Step 4: Energy estimate.
We turn now to the derivation of Lemma 6.2. Set $R_N := \varphi_N - \Psi_N$, then

$$\begin{aligned}\partial_\tau \int_{\mathbf{T}^d\times\mathbf{R}^d} R_N^2\ dy M_\beta(v)dv = &-\frac{\lambda_N}{2}\int_{\mathbf{T}^d\times\mathbf{R}^d} v\cdot\nabla_y\left(R_N^2\right)\ dy M_\beta(v)dv \\ &- \lambda_N^2 \int_{\mathbf{T}^d\times\mathbf{R}^d} R_N\mathcal{L}R_N dy M_\beta(v)dv + \int_{\mathbf{T}^d\times\mathbf{R}^d} R_N\ S_N\ dy M_\beta(v)dv\,.\end{aligned}$$

The first term on the RHS is equal to zero thanks to the periodic boundary conditions, the second term is non-positive (see Lemma 6.1). Thus

$$\partial_\tau \int_{\mathbf{T}^d\times\mathbf{R}^d} R_N^2\ dy M_\beta(v)dv \leq \left(\int_{\mathbf{T}^d\times\mathbf{R}^d} R_N^2\ dy M_\beta(v)dv\right)^{1/2}\ \left(\int_{\mathbf{T}^d\times\mathbf{R}^d} S_N^2\ dy M_\beta(v)dv\right)^{1/2}\,.$$

It remains to check that

$$\lim_{N\to\infty}\int_{\mathbf{T}^d\times\mathbf{R}^d} S_N^2 \; dyM_\beta(v)dv = 0\,, \tag{6.14}$$

$$\lim_{N\to\infty}\int_{\mathbf{T}^d\times\mathbf{R}^d} dyM_\beta(v)dv \;\left(\varphi_N(\tau=0)-\Psi_N(\tau=0)\right)^2 = 0\,. \tag{6.15}$$

Since b and $D_{k,\ell}$ are in $L^2(\mathbf{R}^d, aM_\beta dv)$, it is enough to consider only the terms depending on y. First note that S_N involves spatial derivatives of u_N of order at most 4, thus from the maximum principle, each term of S_N is bounded in L^∞ norm by $\lambda_N^{(4+d)\zeta-1}$. Since $\zeta<\frac{1}{4d}$, the integral (6.14) vanishes.
Concerning the initial data (6.15), the leading contributions of φ_N and u_N cancel and the corrections tend to 0 when N goes to infinity. Thus (6.15) holds.
Lemma 6.2 follows then by applying a Gronwall Lemma. □

6.2. **Convergence to the brownian motion.** To prove the convergence of the rescaled tagged particle Ξ to a brownian motion, one needs to check (see [5])

- the convergence of the marginals of the tagged particle sampled at different times

$$\lim_{N\to\infty}\mathbb{E}\Big(h_1\big(\Xi(\tau_1)\big)\dots h_\ell\big(\Xi(\tau_\ell)\big)\Big) = \mathbb{E}\Big(h_1\big(B(\tau_1)\big)\dots h_\ell\big(B(\tau_\ell)\big)\Big)\,. \tag{6.16}$$

  We stress that these marginals refer to time averages and not to the number of particles.
- the tightness of the sequence, i.e. that is for any $\tau\in[0,T]$

$$\forall \xi>0, \qquad \lim_{\eta\to 0}\lim_{N\to\infty}\mathbb{P}\left(\sup_{\tau<s<\tau+\eta}\big|\Xi(s)-\Xi(\tau)\big|\geq\xi\right)=0\,. \tag{6.17}$$

Step 1. Finite dimensional marginals

First, we are going to rewrite the time correlations in terms of collision trees. Let $t_1<\dots<t_\ell$ be an increasing collection of times and $H_\ell=\{h_i\}_{i\leq\ell}$ a collection of ℓ smooth functions. Define the biased distribution at time $t>t_\ell$ as follows:

$$\begin{aligned}\int_{\mathbf{T}^{Nd}_{\lambda_N}\times\mathbf{R}^{Nd}} dZ_N f_{N,H_\ell}(t,Z_N)\,\Phi(Z_N) &= \mathbb{E}\Big(h_1\big(x_1(t_1)\big)\dots h_\ell\big(x_1(t_\ell)\big)\,\Phi\big(Z_N(t)\big)\Big)\,,\\ &= \int_{\mathbf{T}^{Nd}_{\lambda_N}\times\mathbf{R}^{Nd}} f_N^0(Z_N)\; h_1\big(x_1(t_1)\big)\dots h_\ell\big(x_1(t_\ell)\big)\,\Phi\big(Z_N(t)\big)\,,\end{aligned}$$

for any test function Φ. We stress that by construction

- the biased distribution is in general no longer normalized by 1.
- the measure $f_{N,H_\ell}(t,Z_N)$ is symmetric with respect to the $N-1$ last variables.

The corresponding marginals are

$$f^{(s)}_{N,H_\ell}(t,Z_s) = \int f_{N,H_\ell}(t,Z_N)\,dz_{s+1}\dots dz_N\,. \tag{6.18}$$

By construction f_{N,H_ℓ} satisfies the Liouville equation for $t>t_\ell$ and the marginals $f^{(s)}_{N,H_\ell}$ obey the BBGKY hierarchy (4.1) for $s<N$. Applying the iterated Duhamel formula (4.5), we get

$$f^{(s)}_{N,H_\ell}(t) = \sum_{m=0}^{N-s} Q_{s,s+m}(t-t_\ell) f^{(s+m)}_{N,H_\ell}(t_\ell)\,. \tag{6.19}$$

By construction $f_{N,H_\ell}(t_\ell,Z_N) = f_{N,H_{\ell-1}}(t_\ell,Z_N)h_\ell(z_1)$, where the new distribution is now modified by the the first $\ell-1$ functions. This procedure can be iterated up to the initial time. The backward dynamics can be understood in terms of collision trees which are now weighted by the factor $h_1\big(x_1(t_1)\big)\dots h_\ell\big(x_1(t_\ell)\big)$ associated to the motion of the tagged particle

$$\begin{aligned} f^{(1)}_{N,H_\ell}(t) = \sum_{m_1+\dots+m_\ell=0}^{N-1} & Q_{1,1+m_1}(t-t_\ell)\Big(h_\ell Q_{1+m_1,1+m_2}(t_\ell-t_{\ell-1})\Big(h_{\ell-1}\dots \\ & Q_{1+m_1+\dots+m_{\ell-1},1+m_1+\dots+m_\ell}(t_1)\Big) f_N^{(m_1+\dots+m_\ell)}(0)\,. \end{aligned} \tag{6.20}$$

We stress the fact that this identity holds for any N and any time (independently of the rescaling in λ_N^2).

In order to check (6.17), we need also to generalize the identity to consider correlations of the form

$$\mathbb{E}\Big(h\big(x_1(t_1)-x_1(\tau)\big)\dots h\big(x_1(t_\ell)-x_1(\tau)\big)\Big) \tag{6.21}$$

for a smooth function h with $\tau<t_1<\dots<t_\ell$. Using a partition of unity $\{\Gamma_i^\xi\}$ centered at points $\gamma_i\in\mathbf{T}^d$ with mesh ξ, one can approximate h

$$h(x-y)=\sum_{i,j} h(\gamma_i-\gamma_j)\Gamma_j^\xi(x)\Gamma_i^\xi(y)+O(\xi)\,.$$

This allows us to use the identity (6.20) for any accuracy $\xi>0$ of the approximation. Thus (6.21) can be computed in terms of collision trees which are now weighted by the factor $h\big(x_1(t_1)-x_1(\tau)\big)\dots h\big(x_1(t_\ell)-x_1(\tau)\big)$.

As in (4.8), one can define the limiting distribution which obeys equations similar to (6.19)

$$g^{(s)}_{N,H_\ell}(t)=\sum_{m\geq 0} Q^0_{s,s+m}(t-t_\ell)g^{(s+m)}_{N,H_\ell}(t_\ell)\,.$$

As in (4.9), one has

$$g^{(s)}_{N,H_\ell}(t,Z_s)=g_{N,H_\ell}(t,z_1)\prod_{i=2}^s M_\beta(v_i)\,, \tag{6.22}$$

where $g_{N,H_\ell}(t)$ is the modified density of a tracer $x_1^0(t)$ satisfying the linear Boltzmann dynamics (6.1)

$$\int_{\mathbf{T}^d_{\lambda_N}\times\mathbf{R}^d} dx\; g_{N,H_\ell}(t,x)\,\Phi(x)=\mathbb{E}\Big(h_1\big(x_1^0(t_1)\big)\dots h_\ell\big(x_1^0(t_\ell)\big)\,\Phi\big(x_1^0(t)\big)\Big)\,.$$

Step 2. Approximation of the finite dimensional marginals
Suppose now that the collection H_ℓ satisfies the uniform bounds on $\mathbf{T}^d_{\lambda_N}$

$$\forall i\leq k,\qquad 0\leq h_i(x_1)\leq\alpha,\qquad \big|\nabla_{x_1}h_i(x_1)\big|\leq\alpha'\,. \tag{6.23}$$

Thus the $f^{(s)}_{N,H_\ell}$ satisfy the maximum principle (3.5) with an extra factor α^k. Note that a lower bound on the functions $\{h_i\}$ is not needed as $f^{(s)}_{N,H_\ell}$ is not normalized.
As the maximum principle holds, the pruning procedure on the collision trees applies also in this case and allows us to restrict to trees with at most 2^K collisions during the time interval $[0,t]$. Furthermore, the comparison of the trajectories for $f^{(1)}_{N,H_\ell}$ and g_{N,H_ℓ} can be achieved in the same way as before on a tree with less than 2^K collisions and no recollisions. Analogous bounds as in Proposition 5.7 can be obtained, but one has to take into account that the trees are now weighted by $h_1\big(\frac{1}{\lambda_N}x_1^0(\lambda_N^2\,\tau_1)\big)\dots h_\ell\big(\frac{1}{\lambda_N}x_1^0(\lambda_N^2\,\tau_\ell)\big)$. Using Lemma 5.6, one can restrict to trajectories such that $\big|x_1^0(t)-x_1(t)\big|\leq 2^K\varepsilon$ for any time $t\in[0,\lambda_N^2 T]$.

These small deviations in the trajectories lead to a correction of the weight of order $\frac{\alpha'}{\lambda_N}\ell$. Thus, bounds similar to (6.3) holds for any $\tau \in [0,T]$

$$\left\| \lambda_N^d f_{N,H_\ell}^{(1)}(\lambda_N^2 \tau, \lambda_N y, v) - g_{N,H_\ell}(\lambda_N^2 \tau, \lambda_N y, v) \right\|_{L^\infty([0,T]\times \mathbf{T}^d \times \mathbf{R}^d)} \leq C \frac{T^4}{\big(\log\log N\big)^{1/10}} \,. \tag{6.24}$$

The position of the tagged particle $x_1^0(t) = x_1^0(0) + \int_0^t v_1^0(s)\, ds$ is an additive functional of the Markov chain $\{v_1^0(s)\}_{s\geq 0}$ with generator $\mathcal{L}$ introduced in (6.7). As the generator $\mathcal{L}$ has a spectral gap, the invariance principle holds for the tracer x_1^0 (see [21] theorem 2.32 page 74). This implies the convergence of the rescaled finite dimensional marginals towards the ones of the brownian motion B with variance κ, i.e. that for any smooth functions $\{h_i\}_{i\leq \ell}$ taking values in $\mathbf{T}^d$

$$\lim_{N\to\infty} \mathbb{E}\Big(h_1\big(\frac{1}{\lambda_N} x_1^0(\lambda_N^2\, \tau_1)\big) \dots h_\ell\big(\frac{1}{\lambda_N} x_1^0(\lambda_N^2\, \tau_\ell)\big)\Big) = \mathbb{E}\Big(h_1\big(B(\tau_1)\big) \dots h_\ell\big(B(\tau_\ell)\big)\Big) \,. \tag{6.25}$$

Note that the diffusion coefficient κ defined in (6.10) coincides with the probabilist interpretation (see [21] page 47)

$$\kappa = \lim_{t\to\infty} \mathbb{E}_{M_\beta}\left[\left(\frac{1}{\sqrt{t}} \int_0^t v_1^0(s)\, ds \right)^2 \right] .$$

Combining (6.24) and (6.25), we deduce the convergence of the finite dimensional marginals (6.16).

Step 3. Tightness

In order to evaluate (6.17), it is enough to sample the trajectory of the tagged particle at the times $\tau_i = \{\tau + \frac{u}{\lambda_N^2} i\}_{i\leq \ell_N}$ for some small $u>0$ (to be tuned later) and with $\ell_N = \lambda_N^2\, \eta/u$. We can decompose the path deviations into two terms

$$\begin{aligned} \mathbb{P}\left(\sup_{\tau<s<\tau+\eta} \big|\Xi(s) - \Xi(\tau)\big| \geq 2\xi \right) \leq & 1 - \mathbb{P}\left(\bigcap_{i=1}^{\ell_N} \big\{ \big|\Xi(\tau_i) - \Xi(\tau)\big| < \xi \big\} \right) \\ & + \sum_{i=1}^{\ell_N} \mathbb{P}\left(\sup_{\tau_i<s<\tau_{i+1}} \big|\Xi(s) - \Xi(\tau_i)\big| \geq \xi \right) . \end{aligned} \tag{6.26}$$

We shall first evaluate the last term in the RHS which involves only events occurring in a microscopic time scale of length u. Given $i\leq \ell_N$, let $t_i = iu + \lambda_N^2 \tau$ and $t_{i+1} = u + t_i$ then

$$\mathbb{P}\left(\sup_{\tau_i<s<\tau_{i+1}} \big|\Xi(s) - \Xi(\tau_i)\big| \geq \xi \right) = \mathbb{P}\left(\sup_{t_i<s<t_{i+1}} \big|x_1(s) - x_1(t_i)\big| \geq \lambda_N \xi \right) .$$

In order to control the tagged particle fluctuations during the time length of order u, it is enough to bound its velocity in the time interval $[t_i, t_{i+1}]$ as it has been done in Proposition 5.1. By choosing $u\in(0,1)$ small enough (depending only on β and d, but not on N), the iterated Duhamel formula (4.5) is converging uniformly up to time u thanks to the estimate of Lemma 4.1 and the maximum principle applied at time t_i. Thus cutting-off the initial high kinetic energies, one gets

$$\|\lambda^d (f_N^{(1)} - f_{N,E}^{(1)})\|_{L^\infty([t_i,t_{i+1}]\times \mathbf{T}_\lambda^d \times \mathbf{R}^d)} \leq \mu_N^2 e^{-\frac{\beta}{4}E^2} \,, \tag{6.27}$$

with notation similar to those in Proposition 5.1. For $E = \sqrt{\lambda_N}$, this term tends to 0 as N goes to infinity. This shows that one can restrict the data at time t_i to configurations with a kinetic energy bounded by λ_N. Since the kinetic energy is conserved by the elastic collisions, one deduces that the velocity of the tagged particle is uniformly bounded by $\sqrt{\lambda_N}$ in $[t_i, t_{i+1}]$. Thus after energy truncation, the tagged particle satisfies

$$\sup_{t_i \leq s \leq t_i+1} \big|x_1(s) - x_1(t_i)\big| \leq u\sqrt{\lambda_N} \ll \xi\lambda_N \,.$$

Repeating estimate (6.27) in each small time interval and using that $\mu_N \leq \lambda_N^{1/4}$ (see (6.2)), one gets

$$\sum_{i=1}^{\ell_N} \mathbb{P}\left(\sup_{t_i \leq s \leq t_{i+1}} \big|x_1(s) - x_1(t_i)\big| \geq \xi\lambda_N\right) \leq \lambda_N^{5/2}\frac{\eta}{u}e^{-\frac{\beta}{4}\lambda_N} \,.$$

This tends to 0 when λ_N tends to infinity uniformly in $\xi < 1$.

To estimate the first term in (6.26), we introduce $h_{N,\xi}$ a smooth function such that $1_{\{|z| \leq \lambda_N\xi/2\}} \leq h_{N,\xi}(z) \leq 1_{\{|z| \leq \lambda_N \xi\}}$. One can choose $h_{N,\xi}$ such that

$$0 \leq h_{N,\xi}(z) \leq 1, \qquad \big|\nabla_z h_{N,\xi}(z)\big| \leq \frac{2}{\lambda_N \xi} \,. \tag{6.28}$$

Since

$$\mathbb{E}\big(\prod_{i=1}^{\ell_N} 1_{\left|\Xi(\tau_i) - \Xi(\tau)\right| \leq \xi}\big) \geq \mathbb{E}\big(\prod_{i=1}^{\ell_N} h_{N,\xi}\big(x_1(\lambda_N^2\tau_i) - x_1(\lambda_N^2\tau)\big)\big) \,,$$

it is enough to prove that the RHS converges to 1. At this stage, we are going to use that probabilities of the form (6.21) can also be evaluated in terms of weighted trees as in step 1. Since $h_{N,\xi}$ is bounded by 1, the maximum principle applies uniformly in ℓ_N. The tree decomposition and the reduction to non pathological trajectories hold as in the previous proof. In the approximation of the tagged particle x_1 by x_1^0, one has to modify the bounds of Proposition 5.7. Indeed, the trees are now weighted by $\prod_{i=1}^{\ell_N} h_{N,\xi}\big(x_1(\lambda_N^2\tau_i) - x_1(\lambda_N^2\tau)\big)$ and $\prod_{i=1}^{\ell_N} h_{N,\xi}\big(x_1^0(\lambda_N^2\tau_i) - x_1^0(\lambda_N^2\tau)\big)$. Using Lemma 5.6, one can restrict to trajectories such that $\big|x_1^0(t) - x_1(t)\big| \leq 2^K\varepsilon$ for any time $t \in [0, \lambda_N^2 T]$. Similar computation as in Proposition 5.7 leads to an upper bound with an extra factor $\ell_N\big\|\nabla_z h_{N,\xi}\big\|_\infty$ which can be controled by (6.28):

$$\begin{aligned}\Big\|\widetilde{g}^{(1,K)}_{N,H_{\ell_N},E,\delta} - \widetilde{g}^{(1,K)}_{N,H_{\ell_N},E,\delta,\varepsilon}\Big\|_{L^\infty([0,t]\times\mathbf{T}^d_\lambda\times\mathbf{R}^d)} &\leq C\left(\Lambda_N\mu_N + \frac{2\ell_N}{\xi\lambda_N}\right) \\ &\quad \times \mu_N^2 2^{K(K+1)}(Ct)^{2^{K+1}}\, 2^K\varepsilon \,,\end{aligned} \tag{6.29}$$

where $\widetilde{g}^{(K)}_{N,H_{\ell_N},E,\delta,\varepsilon}$ stands now for the truncated density weighted by the functions H_{ℓ_N}. Thus a bound similar to (6.24) holds

$$\begin{aligned}\Big|\mathbb{E}\big(\prod_{i=1}^{\ell_N} h_{N,\xi}\big(x_1(\lambda_N^2\tau_i) - x_1(\lambda_N^2\tau)\big)\big) - \mathbb{E}\big(\prod_{i=1}^{\ell_N} h_{N,\xi}\big(x_1^0(\lambda_N^2\tau_i) - x_1^0(\lambda_N^2\tau)\big)\big)\Big| \\ \leq C\frac{T^4}{\big(\log\log N\big)^{1/10}} \,.\end{aligned} \tag{6.30}$$

The tightness for the process x_1^0 derived in [21] (Theorem 2.32 page 74) implies that for any $\xi > 0$ and $\ell_N = \lambda_N^2\eta/u$

$$\lim_{\eta\to 0} \mathbb{P}\left(\bigcap_{i=1}^{\ell_N} \big\{\big|x_1^0(\lambda_N^2\tau_i) - x_1^0(\lambda_N^2\tau)\big| < \xi\lambda_N/2\ \big\}\right) = 1 \,.$$

By the trajectory approximation (6.30), the same property holds as well for x_1. Using (6.26), this completes the proof of (6.17).

Appendix A. Asymptotic control of the exclusion

For the sake of completeness, we recall here the proof of Proposition 2.1. We omit all subscripts β to simplify the presentation.

• *First step: asymptotic behaviour of the partition function.*

We first prove that in the scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$,

$$1 \leq \lambda^{ds}\bar{\mathcal{Z}}_N^{-1}\bar{\mathcal{Z}}_{N-s} \leq \big(1-\varepsilon\kappa_d\big)^{-s}, \tag{A.1}$$

where κ_d denotes the volume of the unit ball in $\mathbf{R}^d$. The first inequality is due to the immediate upper bound

$$\bar{\mathcal{Z}}_N \leq \lambda^{ds}\bar{\mathcal{Z}}_{N-s}\,.$$

Let us prove the second inequality. We have by definition

$$\bar{\mathcal{Z}}_{s+1} = \int_{\mathbf{T}_\lambda^{d(s+1)}} \Big(\prod_{1\leq i\neq j\leq s+1} \mathbf{1}_{|x_i-x_j|>\varepsilon}\Big)\, dX_{s+1}\,.$$

By Fubini's equality, we deduce

$$\bar{\mathcal{Z}}_{s+1} = \int_{\mathbf{T}_\lambda^{ds}} \left(\int_{\mathbf{T}_\lambda^{d}} \Big(\prod_{1\leq i\leq s} \mathbf{1}_{|x_i-x_{s+1}|>\varepsilon}\Big)\, dx_{s+1}\right) \Big(\prod_{1\leq i\neq j\leq s} \mathbf{1}_{|x_i-x_j|>\varepsilon}\Big)\, dX_{s}\,.$$

Since

$$\int_{\mathbf{T}_\lambda^{d}} \Big(\prod_{1\leq i\leq s} \mathbf{1}_{|x_i-x_{s+1}|>\varepsilon}\Big)\, dx_{s+1} \geq \lambda^d - \kappa_d s\varepsilon^d\,,$$

we deduce the lower bound

$$\bar{\mathcal{Z}}_{s+1} \geq \bar{\mathcal{Z}}_s(\lambda^d-\kappa_d s\varepsilon^d) \geq \bar{\mathcal{Z}}_s\lambda^d(1-\kappa_d\varepsilon)\,,$$

where we used $s\leq N$ and the scaling $N\varepsilon^{d-1}\lambda^{-d}\equiv 1$. This implies by induction

$$\bar{\mathcal{Z}}_N \geq \bar{\mathcal{Z}}_{N-s}\lambda^{ds}\big(1-\varepsilon\kappa_d\big)^{s}\,.$$

That proves (A.1).

• *Second step: convergence of the marginals.*

Let us introduce the short-hand notation

$$dZ_{(s+1,N)} := dz_{s+1}\dots dz_N\,.$$

We compute for $s\leq N$:

$$\begin{aligned} M_N^{(s)}(Z_s) = \bar{\mathcal{Z}}_N^{-1}\mathbf{1}_{Z_s\in\mathcal{D}_\varepsilon^s}&\left(\frac{\beta}{2\pi}\right)^{\frac{sd}{2}}\exp\left(-\frac{\beta}{2}|V_s|^2\right)\\ &\int_{\mathbf{R}^{d(N-s)}}\left(\frac{\beta}{2\pi}\right)^{\frac{(N-s)d}{2}}\exp\left(-\frac{\beta}{2}\sum_{i=s+1}^N|v_i|^2\right)dV_{(s+1,N)}\\ &\int_{\mathbf{T}_\lambda^{d(N-s)}}\left(\prod_{s+1\leq i\neq j\leq N}\mathbf{1}_{|x_i-x_j|>\varepsilon}\right)\left(\prod_{i'\leq s<j'}\mathbf{1}_{|x_{i'}-x_{j'}|>\varepsilon}\right)dX_{(s+1,N)}\,. \end{aligned}$$

We deduce, by symmetry,

$$M_N^{(s)} = \bar{\mathcal{Z}}_N^{-1}\mathbf{1}_{Z_s\in\mathcal{D}_\varepsilon^s}M^{\otimes s}\Big(\bar{\mathcal{Z}}_{N-s}-\bar{\mathcal{Z}}^\flat_{(s+1,N)}\Big) \tag{A.2}$$

with the notation

$$\bar{\mathcal{Z}}^\flat_{(s+1,N)} := \int_{\mathbf{T}_\lambda^{d(N-s)}} \Big(1 - \prod_{i\leq s<j} \mathbf{1}_{|x_i-x_j|>\varepsilon}\Big) \prod_{s+1\leq k\neq \ell\leq N} \mathbf{1}_{|x_k-x_\ell|>\varepsilon}\, dX_{(s+1,N)}\,.$$

From there, the difference $\mathbf{1}_{Z_s\in\mathcal{D}^s_\varepsilon} M^{\otimes s} - \lambda^{ds} M_N^{(s)}$ decomposes as a sum:

$$\begin{aligned}\mathbf{1}_{Z_s\in\mathcal{D}^s_\varepsilon} M^{\otimes s} - \lambda^{ds} M_N^{(s)} &= \Big(1 - \lambda^{ds}\bar{\mathcal{Z}}_N^{-1}\bar{\mathcal{Z}}_{N-s}\Big)\mathbf{1}_{Z_s\in\mathcal{D}^s_\varepsilon} M^{\otimes s} \\ &\quad + \lambda^{ds}\bar{\mathcal{Z}}_N^{-1}\bar{\mathcal{Z}}^\flat_{(s+1,N)}\mathbf{1}_{Z_s\in\mathcal{D}^s_\varepsilon} M^{\otimes s}\,.\end{aligned} \tag{A.3}$$

By (A.1), there holds $1 - \lambda^{ds}\bar{\mathcal{Z}}_N^{-1}\bar{\mathcal{Z}}_{N-s} \to 0$ as $N\to\infty$, for fixed s. Since $M^{\otimes s}$ is uniformly bounded, this implies that the first term in the right-hand side of (A.3) tends to 0 as N goes to ∞. Besides, by

$$0 \leq 1 - \prod_{i\leq s<j} \mathbf{1}_{|x_i-x_j|>\varepsilon} \leq \sum_{i\leq s<j} \mathbf{1}_{|x_i-x_j|<\varepsilon}\,,$$

we bound

$$\bar{\mathcal{Z}}^\flat_{(s+1,N)} \leq \sum_{1\leq i\leq s} \int_{\mathbf{T}_\lambda^{d(N-s)}} \Big(\sum_{s+1\leq j\leq N} \mathbf{1}_{|x_i-x_j|<\varepsilon}\Big) \prod_{s+1\leq k\neq \ell\leq N} \mathbf{1}_{|x_k-x_\ell|>\varepsilon}\, dX_{(s+1,N)}\,.$$

Given $1\leq i\leq s$, there holds by symmetry and Fubini's equality,

$$\begin{aligned}&\int_{\mathbf{T}_\lambda^{d(N-s)}} \Big(\sum_{s+1\leq j\leq N} \mathbf{1}_{|x_i-x_j|<\varepsilon}\Big) \prod_{s+1\leq k\neq l\leq N} \mathbf{1}_{|x_k-x_l|>\varepsilon}\; dX_{(s+1,N)} \\ &\leq (N-s)\int_{\mathbf{T}_\lambda^d} \mathbf{1}_{|x_i-x_{s+1}|<\varepsilon}\, dx_{s+1} \int_{\mathbf{T}_\lambda^{d(N-s-1)}} \prod_{s+2\leq k\neq l\leq N} \mathbf{1}_{|x_k-x_l|>\varepsilon}\, dX_{(s+2,N)} \\ &= (N-s)\int_{\mathbf{T}_\lambda^d} \mathbf{1}_{|x_i-x_{s+1}|<\varepsilon}\, dx_{s+1} \;\times\; \bar{\mathcal{Z}}_{N-s-1}\,,\end{aligned}$$

so that

$$\bar{\mathcal{Z}}^\flat_{(s+1,N)} \leq s(N-s)\varepsilon^d\kappa_d\bar{\mathcal{Z}}_{N-s-1}\,. \tag{A.4}$$

By (A.1), we obtain

$$\lambda^{ds}\bar{\mathcal{Z}}_N^{-1}\bar{\mathcal{Z}}^\flat_{(s+1,N)} \leq \varepsilon s\kappa_d\big(1-\varepsilon\kappa_d\big)^{-(s+1)}\,,$$

and the upper bound tends to 0 as $N\to\infty$, for fixed s. This implies convergence to 0 of the second term in the right-hand side of (A.3). Finally since $\mathbf{1}_{\mathcal{D}^s_\varepsilon} \to 1$ in $\mathcal{D}^s_{\varepsilon_0}$, the result follows. □

Appendix B. Recollisions in the torus

We show here how to adapt the arguments of [14] to prove Lemma 5.3.

• To build the set of "bad velocities", we use the correspondence between the torus and the whole space with periodic structure. Asking that there exists $s\in[0,t]$ such that

$$d\big((x_1-v_1s),(x_2-v_2s)\big) \leq \varepsilon\,,$$

boils down to having

$$(x_1-v_1s)-(x_2-v_2s) \in \bigcup_{k\in\mathbf{Z}^d} B_{\lambda k}(\varepsilon)\,.$$

Then, by the triangular inequality and provided that $\varepsilon < \bar a$,

$$(\bar x_1 - v_1 s) - (\bar x_2 - v_2 s) \in \bigcup_{k\in \mathbf{Z}^d} B_{\lambda k}(3\bar a)\,.$$

Now, since $|v_1 - v_2| \leq 2E$ and $s \in [0,t]$, this implies that

$$s(v_1-v_2) \in \Big(\bigcup_{k\in \mathbf{Z}^d} B_{\bar x_1 - \bar x_2 + \lambda k}(3\bar a)\Big) \cap B_{2Et}(0)\,.$$

In other words, $v_1 - v_2$ has to belong to a finite union of cones of vertex 0

- at most one of which is of solid angle $(\bar a/\varepsilon_0)^{d-1}$;
- the other ones (at most $\left(4E\frac{t}{\lambda}\right)^d$) are of solid angle $(12\bar a/\lambda)^{d-1}$.

The intersection $K(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)$ of these cones and of the sphere of radius $2E$ is of size

$$|K(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)| \leq CE^d \left(\left(\frac{\bar a}{\varepsilon_0}\right)^{d-1} + \left(E\frac{t}{\lambda}\right)^d \left(\frac{\bar a}{\lambda}\right)^{d-1}\right)\,.$$

• In order to prove the second estimate, we need to refine a little bit the previous argument. Asking that there exists $s \in [\delta, t]$ such that

$$d((x_1 - v_1 s), (x_2 - v_2 s)) \leq \varepsilon_0\,,$$

boils down to having

$$s(v_1 - v_2) \in B(\bar x_1 - \bar x_2 + \lambda k, 3\varepsilon_0)\,, \tag{B.1}$$

for some $k \in \mathbf{Z}^d \cap B(\bar x_2 - \bar x_1, 2Et/\lambda)$.

- If $|\bar x_1 - \bar x_2 + \lambda k| \geq \lambda/4$, condition (B.1) implies that $v_1 - v_2$ belongs to the intersection of $B(0,2E)$ and some cone of vertex 0 and solid angle $(\varepsilon_0/\lambda)^{d-1}$.
- If $|\bar x_1 - \bar x_2 + \lambda k| \leq \lambda/4$ (which can happen only for one value of k), denoting by n_k any unit vector normal to $\bar x_1 - \bar x_2 + \lambda k$, we deduce from (B.1) that
$$s|(v_1 - v_2)\cdot n| \leq 3\varepsilon_0$$
from which we deduce that $v_1 - v_2$ belongs to the intersection of $B(0,2E)$ and some cylinder of radius ε_0/δ.

The union $K_\delta(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)$ of these "bad" sets is therefore of size

$$|K_\delta(\bar x_1 - \bar x_2, \varepsilon_0, \bar a)| \leq CE \left((\varepsilon_0/\delta)^{d-1} + E^{d-1}\left(E\frac{t}{\lambda}\right)^d \left(\frac{\varepsilon_0}{\lambda}\right)^{d-1}\right)\,.$$

The lemma is proved. □

Acknowledgments. The authors are grateful to Laurent Desvillettes, François Golse, Herbert Spohn and Balint Toth for fruitful discussions and for suggesting interesting bibliographical references.

References

[1] R. Alexander, The Infinite Hard Sphere System, *Ph.D. dissertation,* Dept. Mathematics, Univ. California, Berkeley, 1975.

[2] C. Bardos, F. Golse, C.D. Levermore: *Fluid Dynamic Limits of Kinetic Equations II: Convergence Proofs for the Boltzmann Equation*, Comm. Pure & Appl. Math **46** (1993), 667–753.

[3] C. Bardos, R. Santos, R. Sentis, Diffusion approximation and computation of the critical size. *Trans. Amer. Math. Soc*, **284** (2), p. 617-649 (1984).

[4] Th. Bodineau, I. Gallagher, L. Saint-Raymond, *in preparation.*

[5] P. Billingsley, *Probability and measure*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1995.

[6] C. Boldrighini, L.A. Bunimovich, Y. Sinai, On the Boltzmann equation for the Lorentz gas, *Journal of statistical physics*, **32** (3), p. 477-501 (1983).

[7] E. Caglioti, F. Golse. On the distribution of free path lengths for the periodic Lorentz gas III. *Commun. Math. Phys.* **236** (2003), 199-221.

[8] E. Caglioti, F. Golse. On the Boltzmann-Grad limit for the two dimensional periodic Lorentz gas. *J. Stat. Phys.* **141** (2010), 264-317.

[9] C. Cercignani, R. Illner, M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Springer Verlag, New York NY, 1994.

[10] L. Desvillettes, F. Golse, *A remark concerning the Chapman-Enskog asymptotics*, in "Advances in kinetic theory and computing", B. Perthame ed., 191–203, Ser. Adv. Math. Appl. Sci., 22, World Sci. Publishing, River Edge, NJ, 1994.

[11] L. Desvillettes, M. Pulvirenti. The linear Boltzmann equation for long-range forces: a derivation from particles. *Math. Meth. Mod. Appl. Sci.*

[12] L. Desvillettes, V. Ricci, A rigorous derivation of a linear kinetic equation of Fokker-Planck type in the limit of grazing collisions. *J. Statist. Phys.* **104** (2001), no. 5-6, p. 1173-1189.

[13] R. DiPerna and P.-L. Lions, On the Cauchy problem for Boltzmann equations: global existence and weak stability, *Ann. of Math.* (2) **130** (1989), no. 2, p. 321-366.

[14] I. Gallagher, L. Saint-Raymond, B. Texier. From Newton to Boltzmann : the case of hard-spheres and short-range potentials, to appear in ZLAM.

[15] G. Gallavotti. Rigorous theory of the Boltzmann equation in the Lorentz gas. *Università di Roma, Instituto di Fisica, Nota Interna* **358** (1972).

[16] F. Golse. On the periodic Lorentz gas and the Lorentz kinetic equation. *Ann. Fac. Sci. Toulouse Math.* **17** (2008), 735-749.

[17] F. Golse, L. Saint-Raymond: *The Incompressible Navier-Stokes Limit of the Boltzmann Equation for Hard Cutoff Potentials*, J. Math. Pures Appl. **91**, 508–552 (2009).

[18] D. Hilbert. *Sur les problèmes futurs des mathématiques, Compte-Rendu du 2e Congrès International de Mathématiques, tenu à Paris en 1900.* Gauthier-Villars (1902), 58- 114.

[19] D. Hilbert. Begründung der kinetischen Gastheorie. (German) *Math. Ann.* **72** (1912), no. 4, 562-577.

[20] M. Kac, Probability and related topics in physical sciences, *Amer Mathematical Society*, **1** (1959).

[21] T. Komorowski, C. Landim, S. Olla, *Fluctuations in Markov processes. Time symmetry and martingale approximation*, Grundlehren der Mathematischen Wissenschaften, **345** Springer, Heidelberg, 2012.

[22] O.E. Lanford, *Time evolution of large classical systems*, Lect. Notes in Physics **38**, J. Moser ed., 1–111, Springer Verlag (1975).

[23] J. Lebowitz, H. Spohn, Transport properties of the Lorentz gas: Fourier's law. *J. Statist. Phys.* **19** (1978), 633-654.

[24] J. Lebowitz, H. Spohn, Steady state self-diffusion at low density. *J. Statist. Phys.* **29** (1982), 39-55.

[25] H. Lorentz, Le mouvement des électrons dans les métaux. *Arch. Neerl.* **10** (1905), 336-371.

[26] J. Marklof, Kinetic transport in crystals, Proceedings of the XVI International Congress on Mathematical Physics, Prague 2009, *World Scientific* 2010, p. 162-179.

[27] J. Marklof, A. Strömbergsson, The Boltzmann-Grad limit of the periodic Lorentz gas, *Annals of Mathematics* **174** (2011), p. 225-298.

[28] R. Pettersson, On weak and strong convergence to equilibrium for solutions to the linear Boltzmann equation. *J. Statist. Phys.***72** (1993), 355-380.

[29] M. Pulvirenti, C. Saffirio and S. Simonella, On the validity of the Boltzmann equation for short range potentials `arXiv:1301.2514`.

[30] H. Spohn, The Lorentz process converges to a random flight process. *Commun. Math. Phys.* **60** (1978), 277-290.

[31] H. Spohn, Boltzmann hierarchy and Boltzmann equation, in *Kinetic theories and the Boltzmann equation* (Montecatini, 1981), p. 207-220.

[32] H. Spohn, Large scale dynamics of interacting particles, *Springer-Verlag* **174** (1991).

[33] K. Uchiyama, Derivation of the Boltzmann equation from particle dynamics, *Hiroshima Math. J.* **18** (1988), no. 2, p. 245-297.

[34] S. Ukai, The Boltzmann-Grad Limit and Cauchy-Kovalevskaya Theorem, *Japan J. Indust. Appl. Math.,* **18** (2001), p. 383-392.