The Brownian motion as the limit of a deterministic system of hard-spheres

Thierry Bodineau, Isabelle Gallagher & Laure Saint-Raymond

Inventiones mathematicae

ISSN 0020-9910
Volume 203
Number 2

Invent. math. (2016) 203:493-553
DOI 10.1007/s00222-015-0593-9

Inventiones mathematicae

Vol. 203 Fasc. 2 2016

Springer

Invent. math. (2016) 203:493–553
DOI 10.1007/s00222-015-0593-9

The Brownian motion as the limit of a deterministic system of hard-spheres

Thierry Bodineau[1] · **Isabelle Gallagher**[2] ·
Laure Saint-Raymond[3]

Received: 14 May 2013 / Accepted: 18 March 2015 / Published online: 16 April 2015

Abstract We provide a rigorous derivation of the Brownian motion as the limit of a deterministic system 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 $\alpha = N\varepsilon^{d-1} \to \infty$ (with a suitable diffusive scaling of the observation time). As suggested by Hilbert in his sixth problem, we rely on a kinetic formulation as an intermediate level of description between the microscopic and the fluid descriptions: we use indeed the linear Boltzmann equation to describe one tagged particle in a gas close to global equilibrium. Our proof is based 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

1.1 From microscopic to macroscopic models

We are interested here in describing the macroscopic behavior of a gas consisting of N interacting particles of mass m in a domain $\mathbf{D}$ of $\mathbf{R}^d$, with positions and velocities $(x_i, v_i)_{1\le i\le N} \in (\mathbf{D} \times \mathbf{R}^d)^N$, the dynamics of which is given by

✉ Laure Saint-Raymond
saintray@dma.ens.fr

[1] CMAP, Ecole Polytechnique and CNRS, Palaiseau, France

[2] IMJ, Université Paris Diderot, Paris, France

[3] DMA, Ecole Normale Supérieure and Université Pierre et Marie Curie, Paris, France

$$\frac{dx_i}{dt} = v_i, \qquad m\frac{dv_i}{dt} = -\frac{1}{\varepsilon}\sum_{j\neq i}\nabla\Phi\Big(\frac{x_i - x_j}{\varepsilon}\Big), \tag{1.1}$$

for some compactly supported potential Φ, meaning that the scale for the microscopic interactions is typically ε. We shall actually mainly be interested in the case when the interactions are pointwise (hard-sphere interactions): the presentation of that model is postponed to Sect. 2, see (2.1), (2.2).

In the limit when $N \to \infty$, $\varepsilon \to 0$ with $N\varepsilon^d = O(1)$, it is expected that the distribution of particles averages out to a local equilibrium. The microscopic fluxes in the conservations of empirical density, momentum and energy should therefore converge to some macroscopic fluxes, and we should end up with a macroscopic system of equations (depending on the observation time and length scales). However the complexity of the problem is such that there is no complete derivation of any fluid model starting from the full deterministic Hamiltonian dynamics, regardless of the regime (we refer to [20,36,38] for partial results obtained by adding a small noise in the microscopic dynamics).

For rarefied gases, i.e. under the assumption that there is asymptotically no excluded volume $N\varepsilon^d \ll 1$, Boltzmann introduced an intermediate level of description, referred to as kinetic theory, in which the state of the gas is described by the statistical distribution f of the position and velocity of a typical particle. In the Boltzmann-Grad scaling $\alpha \equiv N\varepsilon^{d-1} = O(1)$, we indeed expect the particles to undergo α collisions per unit time in average and all the correlations to be negligible. Therefore, depending on the initial distribution of positions and velocities in the $2dN$-phase space, the 1-particle density f should satisfy a closed evolution equation where the inverse mean free path α measures the collision rate.

In the fast relaxation limit $\alpha \to \infty$, we then expect the system to relax towards local thermodynamic equilibrium, and the dynamics to be described by some macroscopic equations (depending on the observation time and length scales).

One of the major difficulties to achieve this program (Fig. 1) using kinetic models as an intermediate description is to justify the low density limit $\alpha \equiv N\varepsilon^{d-1}$ on time intervals independent of α. Note that this step is also the most complicated one from the conceptual viewpoint as it should explain the appearance of irreversibility, and dissipation mechanisms.

The best result concerning the low density limit, which is due to Lanford in the case of hard-spheres [28] and King [26] for more general potentials (see also [13,21,44] for a complete proof) is indeed valid only for short times, i.e. breaks down before any relaxation can be observed. The result may indeed be stated as follows [21] (see also [37] for less restrictive assumptions on the potential Φ).

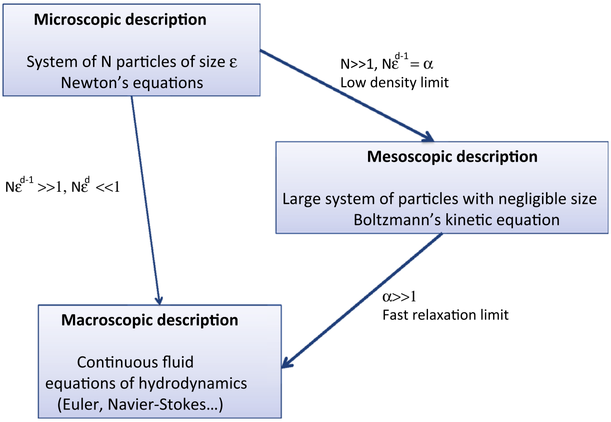

Fig. 1 Fluid equations of hydrodynamics can be recovered directly from the microscopic system or in a two-step limit using Boltzmann's kinetic equation as a mesoscopic description. Note that these two procedures may lead to limiting equations with different transport coefficients since the kinetic equation describes only perfect gases (without excluded volume in the state relation)

Theorem 1.1 *Consider a system of N particles interacting*

- *Either as hard-spheres of diameter ε*
- *Or as in (1.1) via a repulsive potential Φ, with compact support, radial and singular at 0, and such that the scattering of particles can be parametrized by their deflection angle.*

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

$$\left\| f_0 \exp\left(\frac{\beta}{2}|v|^2\right)\right\|_{L^\infty(\mathbf{R}^d_x \times \mathbf{R}^d_v)} \leq \exp(-\mu)$$

for some $\beta > 0$, $\mu \in \mathbf{R}$.

Assume that the N particles are initially distributed according to f_0 and "independent". Then, there exists some $T^* > 0$ (depending only on β and μ) such that, in the Boltzmann-Grad limit $N \to \infty$, $\varepsilon \to 0$, $N\varepsilon^{d-1} = \alpha$, the distribution function of the particles converges uniformly on $[0, T^*/\alpha] \times \mathbf{R}^{2d}$ to the solution of the Boltzmann equation

$$\begin{aligned} \partial_t f + v \cdot \nabla_x f &= \alpha Q(f, f), \\ Q(f, f)(v) &:= \iint_{\mathbf{S}^{d-1} \times \mathbf{R}^d} [f(v^*) f(v_1^*) - f(v) f(v_1)]\, b(v - v_1, \nu)\, dv_1 d\nu \\ v^* &= v + \nu \cdot (v_1 - v)\, \nu, \quad v_1^* = v_1 - \nu \cdot (v_1 - v)\, \nu, \end{aligned} \tag{1.2}$$

with a locally bounded cross-section b depending on Φ 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 that the initial N-particle distribution satisfies a chaos property, namely that the correlations vanish asymptotically. Typically the distribution is obtained by factorization, and conditioning on energy surfaces (see [21] and references therein). In the case of hard-spheres for instance, one would have

$$f_N^0 = \mathcal{Z}_N^{-1} f_0^{\otimes N} \mathbf{1}_{\mathcal{D}_\varepsilon^N} ,$$

with

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

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^0 to 1.

The main difficulty to prove convergence for longer time intervals consists in ruling out the possibility of spatial concentrations of the density leading to some pathological collision process.

1.2 Linear regimes

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

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

$$\begin{aligned} \partial_t \varphi_\alpha + v \cdot \nabla_x \varphi_\alpha &= -\alpha\, \mathcal{L}\varphi_\alpha \\ \mathcal{L}\varphi_\alpha(v) &:= \iint [\varphi_\alpha(v) - \varphi_\alpha(v')]\, M_\beta(v_1)\, b(v - v_1, \nu)\, dv_1 d\nu \\ M_\beta(v) &:= \left(\frac{\beta}{2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{\beta}{2}|v|^2\right), \quad \beta > 0, \end{aligned} \tag{1.3}$$

for times diverging as α when $\alpha \to \infty$. Indeed, in the diffusive regime, the convergence of the Markov process associated to the linear Boltzmann operator $\mathcal{L}$ towards the Brownian motion is by now a classical result [27].

2 Strategy and main results

A 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 initially 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 distribution.
- The obstacles have no dynamics, in particular they do not feel the effect of collisions with the tagged particle.

This problem, suggested by Lorentz [31] 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 [42, Chapter8] for a survey on this topic. We state one basic result due to Gallavotti [22] in the low density limit and then indicate some of the many important research directions.

Theorem 2.1 *Consider randomly distributed scatterers with radius ε in $\mathbf{R}^d$ according to a Poisson distribution of parameter $\alpha\varepsilon^{1-d}$. Let T^t_ε be the flow of a point particle reflected at the boundary of these scatterers. For a given continuous initial datum $f_0 \in L^1 \cap L^\infty(\mathbf{R}^{2d})$, we define*

$$f_\varepsilon(t,x,v) := \mathbb{E}[f_0(T_\varepsilon^{-t}(x,v))].$$

Then, for any time $T>0$, f_ε converges to the solution f of the linear Boltzmann equation (1.3), 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 [41] in terms of convergence of path measures (and not only of the mean density), as well as in

[9] 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 was addressed by Desvillettes, Pulvirenti and Ricci [16,17]. 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. Thus 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.

It is often 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. For the two-dimensional periodic Lorentz gas with fixed scatterer size, Bunimovich and Sinai [10] have shown the convergence, after a suitable time rescaling, of the tagged particle to a Brownian motion. Their method relies on techniques from ergodic theory: it uses the fact that the mapping carrying a phase point on the boundary of a scatterer to the next phase point along its trajectory can be represented by a symbolic dynamics on a countable alphabet which is an ergodic Markov chain on a finite state space. Another important research direction, initiated by Golse, is to consider the periodic Lorentz gas in the Boltzmann-Grad limit. In this case there can be infinitely long free flight paths and the linear Boltzmann equation is no longer a valid limit [12,23,32,33], but the convergence toward a Brownian motion can be recovered after an appropriate superdiffusive rescaling [34].

In [18,19], Erdös, Salmhofer and Yau obtained the counterpart of the long time behavior for random quantum systems. Our approach is closer to their method than to the ones used for the periodic Lorentz gas (even though the setting of [19] deals with a fixed random distribution of obstacles and a slightly different regime, known as weak coupling limit). Their proof is indeed based on a careful analysis of Duhamel's formula in combination with a renormalization of the propagator and stopping rules to control recollisions. We refer also to [14] for further developments on the quantum case.

2.2 Interacting gas of particles

We adopt here a 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. In this paper, we will focus on the case $d \geq 2$ (and refer to [43] for results in the case $d = 1$).

On the one hand, the problem seems more difficult than the Lorentz gas 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 [11,12,23] 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 in 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 4.1). This will be the key to control the collision process, and to prove (like in Kac's model [25] 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 van Beijeren, Lanford, Lebowitz and Spohn [7,30] 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 Lanford's result by considering fluctuations around some global equilibrium. Locally the N-particle distribution f_N should therefore look like a conditioned tensorized Maxwellian.

In the sequel, we shall focus on the case of hard-sphere dynamics (with mass $m = 1$) to avoid technicalities due to artificial boundaries and cluster estimates. We shall further restrict our attention to the case when the domain is periodic $\mathbf{D} = \mathbf{T}^d = [0, 1]^d$ ($d \geq 2$).

The microscopic model is therefore given by the following system of ODEs:

$$\frac{dx_i}{dt} = v_i, \quad \frac{dv_i}{dt} = 0 \quad \text{as long as} \quad |x_i(t) - x_j(t)| > \varepsilon \quad \text{for} \quad 1 \leq i \neq j \leq N, \tag{2.1}$$

with specular reflection after a collision

$$\left.\begin{aligned} v_i(t^+) &= v_i(t^-) - \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j)(t^-) \\ v_j(t^+) &= v_j(t^-) + \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j)(t^-) \end{aligned}\right\} \\ \text{if } |x_i(t) - x_j(t)| = \varepsilon. \tag{2.2}$$

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

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

We now distinguish pre-collisional configurations from post-collisional ones by defining for indexes $1 \le i \neq j \le N$

$$\partial\mathcal{D}_\varepsilon^{N\pm}(i,j) := \Big\{ Z_N \in \mathbf{T}^{dN} \times \mathbf{R}^{dN} \,/\, |x_i - x_j| = \varepsilon, \quad \pm (v_i - v_j)\cdot(x_i - x_j) > 0 \\ \text{and} \quad \forall (k,\ell) \in [1,N]^2 \setminus \{(i,j)\}^2, |x_k - x_\ell| > \varepsilon \Big\}.$$

Given Z_N on $\partial\mathcal{D}_\varepsilon^{N+}(i,j)$, we define $Z_N^* \in \partial\mathcal{D}_\varepsilon^{N-}(i,j)$ as the configuration having the same positions $(x_k)_{1\le k\le N}$, the same velocities $(v_k)_{k\neq i,j}$ for non interacting particles, and the following pre-collisional velocities for particles i and j

$$\begin{aligned} v_i^* &:= v_i - \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j) \\ v_j^* &:= v_j + \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j). \end{aligned}$$

Defining the Hamiltonian

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

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

$$\partial_t f_N + \{H_N, f_N\} = 0 \tag{2.3}$$

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

$$f_N(t, Z_N) = f_N(t, Z_N^*). \tag{2.4}$$

We recall, as shown in [1] for instance, that the set of initial configurations leading to ill-defined characteristics (due to clustering of collision times, or collisions involving more than two particles) is of measure zero in $\mathcal{D}_\varepsilon^N$.

Define the Maxwellian distribution by

$$M_\beta^{\otimes s}(V_s) := \prod_{i=1}^s M_\beta(v_i) \quad \text{and} \quad M_\beta(v) := \left(\frac{\beta}{2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{\beta}{2}|v|^2\right). \tag{2.5}$$

An obvious remark is that M_β is a stationary solution of (1.2), and any function of the energy $f_N \equiv F(H_N)$ is a stationary solution of the Liouville equation (2.3). In particular, for $\beta > 0$, the Gibbs measure with distribution in $\mathbf{T}^{dN} \times \mathbf{R}^{dN}$ defined by

$$\begin{aligned} M_{N,\beta}(Z_N) &:= \frac{1}{\mathcal{Z}_N} \left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(V_N))\, \mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N) \\ &= \frac{1}{\mathcal{Z}_N} \mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N) M_\beta^{\otimes N}(V_N) \end{aligned} \tag{2.6}$$

where the partition function $\mathcal{Z}_N$ is the normalization factor

$$\mathcal{Z}_N := \int_{\mathbf{T}^{dN}\times\mathbf{R}^{dN}} \mathbf{1}_{\mathcal{D}_\varepsilon^N}(Z_N) M_\beta^{\otimes N}(V_N)\, dZ_N = \int_{\mathbf{T}^{dN}} \prod_{1\le i\neq j\le N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\, dX_N, \tag{2.7}$$

and is an invariant measure for the gas dynamics.

In order to obtain the convergence for long times, a natural idea is to "weakly" perturb the equilibrium state $M_{N,\beta}$, by modifying the distribution of one particle. In other words, we shall describe the dynamics of a tagged particle in a background initially at equilibrium. Actually this is the reason for placing the study in a bounded domain, in order for $M_{N,\beta}$ to be integrable in the whole phase space. Moreover we have restricted our attention to the case of a torus in order to avoid pathologies related to boundary effects, and complicated free dynamics.

The strategy of perturbating $M_{N,\beta}$ is classical in probability theory; following this strategy

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

What we shall actually prove is that the limiting dynamics is governed by the linear Boltzmann equation (1.3) with hard-sphere cross-section.

2.3 Main results

For the sake of simplicity, we consider only one tagged particle which will be labeled by 1 with coordinates $z_1 = (x_1, v_1)$. The initial data is a perturbation of the equilibrium density (2.6) only with respect to the position x_1 of the tagged particle. Consider ρ^0 a continuous density of probability on $\mathbf{T}^d$ and define

$$f_N^0(Z_N) := M_{N,\beta}(Z_N)\rho^0(x_1). \tag{2.8}$$

Note that the distribution f_N^0 is normalized by 1 in $L^1(\mathbf{T}^{dN} \times \mathbf{R}^{dN})$ thanks to the translation invariance of $\mathbf{T}^d$ and that $\int_{\mathbf{T}^d} \rho^0(x)dx = 1$.

The main result of our study is the following statement.

Theorem 2.2 *Consider the initial distribution f_N^0 defined in (2.8). Then the distribution $f_N^{(1)}(t, x, v)$ of the tagged particle is close to $M_\beta(v)\varphi_\alpha(t, x, v)$, where $\varphi_\alpha(t, x, v)$ is the solution of the linear Boltzmann equation (1.3) with initial data $\rho^0(x_1)$ and hard-sphere cross section. More precisely, for all $t > 0$ and all $\alpha > 1$, in the limit $N \to \infty$, $N\varepsilon^{d-1}\alpha^{-1} = 1$, one has*

$$\left\| f_N^{(1)}(t, x, v) - M_\beta(v)\varphi_\alpha(t, x, v)\right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} \le C\left[\frac{t\alpha}{(\log\log N)^{\frac{A-1}{A}}}\right]^{\frac{A^2}{A-1}}, \tag{2.9}$$

where $A \ge 2$ can be taken arbitrarily large, and C depends on A, β, d and $\|\rho^0\|_{L^\infty}$.

In [7,30], the linear Boltzmann equation was derived for any time $t > 0$ (independent of N). In comparison, 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. Theorem 2.2 proves that the linear Boltzmann equation is a good asymptotics of the hard-sphere dynamics, even for large concentrations α and long times t. It further provides a rather good estimate on the approximation error. Up to a suitable rescaling of time, we can therefore obtain diffusive limits.

In the macroscopic limit, the trajectory of the tagged particle is defined by

$$\Xi(\tau) := x_1\big(\alpha\tau\big) \in \mathbf{T}^d. \tag{2.10}$$

The distribution of $\Xi(\tau)$ is given by $f_N^{(1)}(\alpha\tau, x, v)$. In the following, τ represents the macroscopic time scale.

Theorem 2.3 *Consider N hard spheres on the space* $\mathbf{T}^d \times \mathbf{R}^d$, *initially distributed according to* f_N^0 *defined in* (2.8). *Assume that* ρ^0 *belongs to* $C^0(\mathbf{T}^d)$. *Then the distribution* $f_N^{(1)}(\alpha\tau, x, v)$ *remains close for the* L^∞*-norm to* $\rho(\tau, x)M_\beta(v)$ *where* $\rho(\tau, x)$ *is the solution of the linear heat equation*

$$\partial_\tau \rho - \kappa_\beta \Delta_x \rho = 0 \quad in \quad \mathbf{T}^d, \quad \rho_{|\tau=0} = \rho^0, \tag{2.11}$$

and the diffusion coefficient κ_β *is given by*

$$\kappa_\beta := \frac{1}{d}\int_{\mathbf{R}^d} v\mathcal{L}^{-1}v\, M_\beta(v)dv,$$

where $\mathcal{L}$ *is the linear Boltzmann operator* (1.3) *and* $\mathcal{L}^{-1}$ *is its pseudo-inverse defined on* $(\mathrm{Ker}\,\mathcal{L})^\perp$ [*see also* (6.8)]. *More precisely,*

$$\left\| f_N^{(1)}(\alpha\tau, x, v) - \rho(\tau, x)M_\beta(v)\right\|_{L^\infty([0,T]\times\mathbf{T}^d\times\mathbf{R}^d)} \to 0 \tag{2.12}$$

in the limit $N \to \infty$, *with* $\alpha = N\varepsilon^{d-1}$ *going to infinity much slower than* $\sqrt{\log\log N}$.

In the same asymptotic regime, the process $\Xi(\tau) = x_1(\alpha\tau)$ *associated with the tagged particle converges in law towards a Brownian motion of variance* κ_β, *initially distributed under the measure* ρ^0.

The Boltzmann-Grad scaling $\alpha = N\varepsilon^{d-1}$ is chosen such that the mean free path is of order $1/\alpha$, i.e. that a particle has on average α collisions per unit time. This explains why in (2.10), the position of the particle is not rescaled. Indeed over a time scale $\alpha\tau$ a particle will encounter $\alpha^2\tau$ collisions which is the correct balance for a diffusive limit. In other words, one can think of α as a parameter tuning the density of the background particles. The positions and velocities are not rescaled with α and are always at the macroscopic scale.

2.4 Generalizations

For the sake of clarity, Theorem 2.3 has been stated in the simplest framework. We mention below several extensions which can be deduced in a straightforward way from the proof of Theorem 2.3.

Several tagged particles: The dynamics of a finite number of tagged particles can be followed and one can show that asymptotically, they converge to independent Brownian motions. This gives an answer to a conjecture raised by Lebowitz and Spohn [29] on the diffusion of colored particles in a fluid.

Interaction potential: Following the arguments in [21,37], the behavior of a tagged particle in a gas with an interaction potential can also be treated.

Initial data: The perturbation on the initial particle could depend on $z_1 = (x_1, v_1)$ instead of depending only on the position x_1. The comparison argument to the linear Boltzmann equation is identical, but the derivation of the diffusive behavior in Sect. 6.1 should be modified to show the relaxation of the velocity to a Maxwellian at the initial stage (see Remark 6.2).

By considering an initial data of the form

$$\rho^0_\alpha(x_1) = \alpha^{d\zeta}\ \rho^0\big(\alpha^\zeta x_1\big) \qquad \text{with} \quad \zeta \ll 1 \tag{2.13}$$

the tagged particle localizes when α goes to infinity. The analysis can be extended to this class of initial data and leads, in the macroscopic limit, to a Brownian motion starting initially from a Dirac mass.

Scalings: We have chosen here to work with macroscopic variables (x, v), i.e. to rescale the particle concentration of the background and to dilate the time with a factor α. However, the diffusive limit can be obtained by many other equivalent scalings involving the space variable. In particular, one could have considered a domain $[0, \lambda]^d$ with a size λ growing and a Boltzmann-Grad scaling $(N/\lambda^d)\varepsilon^{d-1} = 1$. Rescaling space by a factor λ and time by $\lambda^2 \ll \log\log N$ would have led to the same diffusive limit. In fact, one only needs the Knudsen number to be small and of the same order as the Strouhal number [4,39].

2.5 Structure of the paper

Theorem 2.3 is a consequence of Theorem 2.2, as explained in Sect. 6. The core of our study is therefore the proof of Theorem 2.2, which relies on a comparison of the particle system to a limit system known as Boltzmann hierarchy. This hierarchy is obtained formally in Sect. 3 from the hierarchy of equations satisfied by the marginals of f_N, known as the BBGKY hierarchy (which is introduced in Sect. 3). Section 4 is devoted to the control of the branching process that can be associated with the hierarchies, and in particular with the elimination of super-exponential trees; the specificity of the linear framework is crucial in this step, as it makes it possible to compare the solution with the invariant measure globally in time. The actual proof of the convergence of the BBGKY hierarchy towards the Boltzmann hierarchy, on times diverging with N, can be found in Sect. 5.

Some more technical estimates are postponed to Appendix A and B.

3 Formal derivation of the low density limit

Our starting point to study the low density limit is the Liouville equation (2.3) and its projection on the first marginal

$$f_N^{(1)}(t, z_1) := \int f_N(t, Z_N) dz_2 \dots dz_N.$$

Since it does not satisfy a closed equation, we have to consider the whole BBGKY hierarchy (see Paragraph 3.1). 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.

More precisely we shall see that the specific form of the initial data (see Paragraph 3.2) implies that asymptotically we have the following closure

$$\begin{aligned} &f_N^{(2)}(t, z_1, z_2) \\ &:= \int f_N(t, Z_N) dz_3 \dots dz_N \sim f_N^{(1)}(t, z_1) M_\beta(v_2) \sim \varphi_\alpha(t, z_1) M_\beta(v_1) M_\beta(v_2) \end{aligned}$$

where φ_α satisfies the linear Boltzmann equation (1.3) with initial data ρ^0. Thus the limiting hierarchy reduces to the linear Boltzmann equation (see Paragraph 3.3).

3.1 The series expansion

The quantities we shall consider are the marginals

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

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.

A formal computation based on Green's formula leads to the following BBGKY hierarchy for $s < N$

$$\left(\partial_t + \sum_{i=1}^{s} v_i \cdot \nabla_{x_i}\right) f_N^{(s)}(t, Z_s) = \alpha\big(C_{s,s+1} f_N^{(s+1)}\big)(t, Z_s) \tag{3.1}$$

on $\mathcal{D}_\varepsilon^s$, with the boundary condition as in (2.4)

$$f_N^{(s)}(t, Z_s) = f_N^{(s)}(t, Z_s^*) \text{ on } \partial D_\varepsilon^{s+}(i, j).$$

The collision term is defined by

$$\begin{aligned}\big(C_{s,s+1} f_N^{(s+1)}\big)(Z_s) := (N-s)\varepsilon^{d-1}\alpha^{-1}\\ \times\Bigg(\sum_{i=1}^{s}\int_{\mathbf{S}^{d-1}\times\mathbf{R}^d} f_N^{(s+1)}(\dots, x_i, v_i^*, \dots, x_i+\varepsilon\nu, v_{s+1}^*)\\ \times\Big((v_{s+1}-v_i)\cdot\nu\Big)_+ d\nu dv_{s+1}\\ -\sum_{i=1}^{s}\int_{\mathbf{S}^{d-1}\times\mathbf{R}^d} f_N^{(s+1)}(\dots, x_i, v_i, \dots, x_i+\varepsilon\nu, v_{s+1})\\ \times\Big((v_{s+1}-v_i)\cdot\nu\Big)_- d\nu dv_{s+1}\Bigg)\end{aligned} \tag{3.2}$$

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 we used the trace condition on $\partial\mathcal{D}_\varepsilon^N$ to express all quantities in terms of pre-collisional configurations.

The closure for $s = N$ is given by the Liouville equation (2.3). 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 involve 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 labeled by 1 and always belongs to the sub-system under consideration.

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

$$f_N^{(s)}(t) = \sum_{n=0}^{N-s} \alpha^n \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} \mathbf{S}_s(t-t_1) C_{s,s+1} \mathbf{S}_{s+1}(t_1-t_2) C_{s+1,s+2} \dots \mathbf{S}_{s+n}(t_n) f_N^{(s+n)}(0)\, dt_n \dots dt_1, \tag{3.3}$$

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

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

$$Q_{s,s+n}(t) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} \mathbf{S}_s(t-t_1) C_{s,s+1} \mathbf{S}_{s+1}(t_1-t_2) C_{s+1,s+2} \dots \mathbf{S}_{s+n}(t_n) dt_n \dots dt_1 \tag{3.4}$$

so that

$$f_N^{(s)}(t) = \sum_{n=0}^{N-s} \alpha^n Q_{s,s+n}(t) f_N^{(s+n)}(0). \tag{3.5}$$

Remark 3.1 It is not obvious that formula (3.5) makes sense since the transport operator $\mathbf{S}_{s+1}$ is defined only for almost all initial configurations, and the collision operator $C_{s,s+1}$ is defined by some integrals on manifolds of codimension 1. This fact is analyzed in [40] and in the erratum of [21]. In the following, we will rely on the estimates on the collision operator derived in [21].

3.2 Asymptotic factorization of the initial data

The effect of the exclusion in the equilibrium measure vanishes when ε goes to 0 and the particles become asymptotically independent in the following sense.

Proposition 3.2 *Given $\beta > 0$, there is a constant $C > 0$ such that for any fixed $s \geq 1$, the marginal of order s*

$$M_{N,\beta}^{(s)}(Z_s) := \int M_{N,\beta}(Z_N)\, dz_{s+1} \dots dz_N \tag{3.6}$$

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

$$\left| \left(M_{N,\beta}^{(s)} - M_\beta^{\otimes s} \right) \mathbf{1}_{\mathcal{D}_\varepsilon^s} \right| \leq C^s\, \varepsilon\alpha\, M_\beta^{\otimes s} \tag{3.7}$$

where the Maxwellian distribution $M_\beta^{\otimes s}$ was introduced in (2.5).

The proof of Proposition 3.2, by now classical, is recalled in Appendix A for the sake of completeness.

As a consequence of Proposition 3.2, the initial data is asymptotically close to a product measure: the following result is a direct corollary of Proposition 3.2.

Proposition 3.3 *For the initial data f_N^0 given in (2.8), define the marginal of order s*

$$f_N^{0(s)}(Z_s) := \int f_N^0(Z_N)\,dz_{s+1}\dots dz_N = \rho^0(x_1)M_{N,\beta}^{(s)}(Z_s).$$

There is a constant $C > 0$ such that as $N \to \infty$ in the scaling $N\varepsilon^{d-1} = \alpha \ll \frac{1}{\varepsilon}$

$$\left|\left(f_N^{0(s)} - g^{0(s)}\right)\mathbf{1}_{\mathcal{D}_\varepsilon^s}\right| \le C^s\varepsilon\alpha M_\beta^{\otimes s}\|\rho^0\|_{L^\infty},$$

where $g^{0(s)}$ is defined by

$$g^{0(s)}(Z_s) := \rho^0(x_1)M_\beta^{\otimes s}(V_s). \tag{3.8}$$

3.3 The limiting hierarchy and the linear Boltzmann equation

To obtain the Boltzmann hierarchy we start with the expansion (3.5) and compute the formal limit of the collision operator $Q_{s,s+n}$ when ε goes to 0. Recalling that $(N-s)\varepsilon^{d-1}\alpha^{-1} \sim 1$, it is given by

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

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

$$\begin{aligned}(C_{s,s+1}^0 g^{(s+1)})(Z_s) &:= \sum_{i=1}^s\int g^{(s+1)}(\dots,x_i,v_i^*,\dots,x_i,v_{s+1}^*)\Big((v_{s+1}-v_i)\cdot\nu\Big)_+ d\nu dv_{s+1}\\ &\quad-\sum_{i=1}^s\int g^{(s+1)}(\dots,x_i,v_i,\dots,x_i,v_{s+1})\Big((v_{s+1}-v_i)\cdot\nu\Big)_- d\nu dv_{s+1}.\end{aligned} \tag{3.9}$$

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

$$\forall s \geq 1, \quad g_\alpha^{(s)}(t) = \sum_{n\geq 0} \alpha^n Q^0_{s,s+n}(t) g^{0(s+n)}. \tag{3.10}$$

Remark 3.4 In the Boltzmann hierarchy, the collision operators are defined by integrals on manifolds of codimension d, so we shall require that the functions $\left(g_\alpha^{(s)}\right)_{s\geq 1}$ are continuous, which is possible since free transport preserves continuity on $\mathbf{T}^d \times \mathbf{R}^d$.

Consider the initial data (3.8). Then the family $(g_\alpha^{(s)})_{s\geq 1}$ defined by

$$g_\alpha^{(s)}(t, Z_s) := \varphi_\alpha(t, z_1) M_\beta^{\otimes s}(V_s) \tag{3.11}$$

is a solution to the Boltzmann hierarchy with initial data $g^{0(s)}$ since φ_α satisfies the linear Boltzmann equation (1.3) with initial data ρ^0.

We insist that the $g_\alpha^{(s)}$ are not defined as the marginals of some N-particle density.

Remark 3.5 Note that the estimates established in the next section imply actually that $(g_\alpha^{(s)})_{s\geq 1}$ is the unique solution to the Boltzmann hierarchy (see [21]).

Furthermore the maximum principle for the linear Boltzmann equation leads to the following estimate

$$\sup_{t\geq 0} \varphi_\alpha(t, z_1) \leq \|\rho^0\|_{L^\infty}.$$

In the following for the sake of simplicity we write $g_\alpha := g_\alpha^{(1)}$.

4 Control of the branching process

The restriction on the time of validity T^*/α of Lanford's convergence proof (determined by a weighted norm of the initial data) is based on the elimination of "pathological" collision trees, defined by a too large number of branches created in the time interval $[0, T^*/\alpha]$ (typically greater than $n_\varepsilon = O(|\log \varepsilon|)$, see [21] 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.

4.1 A priori estimates coming from the maximum principle

For initial data as (2.8), uniform a priori bounds can be obtained using only the maximum principle for the Liouville equation (2.3).

Proposition 4.1 *For any fixed N, denote by f_N the solution to the Liouville equation* (2.3) *with initial data* (2.8)*, and by $f_N^{(s)}$ its marginal of order s*

$$f_N^{(s)}(t, Z_s) := \int f_N(t, Z_N)\, dz_{s+1} \dots dz_N. \tag{4.1}$$

Then, for any $s \geq 1$, the following bounds hold uniformly with respect to time

$$\sup_t f_N^{(s)}(t, Z_s) \leq M_{N,\beta}^{(s)}(Z_s)\|\rho^0\|_{L^\infty} \leq C^s M_\beta^{\otimes s}(V_s)\|\rho^0\|_{L^\infty}, \tag{4.2}$$

for some $C > 0$, provided that $\alpha\varepsilon \ll 1$.

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 (4.2) does not see this asymmetry.

Proof One has immediately from (2.8) that

$$f_N^0(Z_N) = M_{N,\beta}(Z_N)\rho^0(x_1) \leq M_{N,\beta}(Z_N)\|\rho^0\|_{L^\infty}.$$

Since the maximum principle holds for the Liouville equation (2.3), and as the Gibbs measure $M_{N,\beta}$ is a stationary solution, we get for all $t \geq 0$

$$f_N(t, Z_N) \leq M_{N,\beta}(Z_N)\|\rho^0\|_{L^\infty}.$$

The inequalities for the marginals follow by integration and Proposition 3.2. □

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. In the following we shall denote by $|Q|_{s,s+n}$ the operator obtained by summing the absolute values of all elementary contributions

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

and similarly for $|Q^0|_{s,s+n}$

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

where

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

and

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

For $\lambda > 0$ and $k \in \mathbf{N}^*$, we define $X_{\varepsilon,k,\lambda}$ the space of measurable functions f_k defined almost everywhere on $\mathcal{D}^k_\varepsilon$ such that

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

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

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

Lemma 4.2 *There is a constant C_d depending only on d such that for all $s,n\in\mathbf{N}^*$ and all $t\geq 0$, 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 $X_{\varepsilon,s+n,\lambda}$, $|Q|_{s,s+n}(t)f_{s+n}$ belongs to $X_{\varepsilon,s,\frac{\lambda}{2}}$ and*

$$\Big\||Q|_{s,s+n}(t)f_{s+n}\Big\|_{\varepsilon,s,\frac{\lambda}{2}} \leq e^{s-1}\left(\frac{C_d t}{\lambda^{\frac{d+1}{2}}}\right)^n \|f_{s+n}\|_{\varepsilon,s+n,\lambda}. \tag{4.3}$$

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

$$\left\| |Q^0|_{s,s+n}(t)g_{s+n} \right\|_{0,s,\frac{\lambda}{2}} \le e^{s-1} \left(\frac{C_d t}{\lambda^{\frac{d+1}{2}}} \right)^n \|g_{s+n}\|_{0,s+n,\lambda}. \tag{4.4}$$

Proof Estimate (4.3) is simply obtained from the fact that the transport operators preserve the weighted norms, along with the continuity of the elementary collision operators. From the erratum of [21], we get the following statements

- The transport operators satisfy the identities

$$\|\mathbf{S}_k(t) f_k\|_{\varepsilon,k,\lambda} = \|f_k\|_{\varepsilon,k,\lambda}$$
$$\|\mathbf{S}_k^0(t) g_k\|_{0,k,\lambda} = \|g_k\|_{0,k,\lambda}.$$

- The collision operators satisfy the following bounds in the Boltzmann-Grad scaling $N\varepsilon^{d-1} \equiv \alpha$

$$\left| \mathbf{S}_k(-t)\, |C_{k,k+1}|\, \mathbf{S}_{k+1}(t) f_{k+1}(Z_k) \right| \le C_d\, \lambda^{-\frac{d}{2}} \left(k\lambda^{-\frac{1}{2}} + \sum_{1\le i\le k} |v_i| \right) \times \exp\left(-\lambda H_k(Z_k)\right) \|f_{k+1}\|_{\varepsilon,k+1,\lambda}$$

almost everywhere on $\mathbf{R}_t \times \mathcal{D}_\varepsilon^k$, for some $C_d > 0$ depending only on d, and

$$\left| |C^0_{k,k+1}|\, g_{k+1}(Z_k) \right| \le C_d\, \lambda^{-\frac{d}{2}} \left(k\lambda^{-\frac{1}{2}} + \sum_{1\le i\le k} |v_i| \right) \times \exp\left(-\lambda H_k(Z_k)\right) \|g_{k+1}\|_{0,k+1,\lambda}, \tag{4.5}$$

on $\mathbf{T}^{dk} \times \mathbf{R}^{dk}$.

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

$$\begin{aligned} &\sum_{1\le i\le k} |v_i| \exp\left(-\frac{\lambda}{4n} \sum_{1\le j\le k} |v_j|^2 \right) \\ &\quad \le \left(k\frac{2n}{\lambda} \right)^{\frac{1}{2}} \left(\sum_{1\le i\le k} \frac{\lambda}{2n} |v_i|^2 \exp\left(-\frac{\lambda}{2n} \sum_{1\le j\le k} |v_j|^2 \right) \right)^{1/2} \\ &\quad \le \left(\frac{2nk}{e\lambda} \right)^{1/2} \le \sqrt{\frac{2}{e\lambda}}(s+n), \end{aligned}$$

with $k \leq s + n$ in the last inequality. Each collision operator gives therefore a loss of $C\lambda^{-(d+1)/2}(s+n)$ 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

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

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

4.3 Collision trees of controlled size

For general initial data (in particular, for chaotic initial data), the proof of Lanford's convergence result then relies on two steps:

(i) A short time bound for the series expansion (3.5) 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 of each term of the series.

However after a short time (depending on the initial data), the question of the convergence of the series (3.5) is still open. One of the difficulties to prove this convergence is to take into account the cancellations between the gain and loss terms of the collision operators. These cancellations are neglected in Lanford's strategy.

Here we assume that the BBGKY initial data takes the form (2.8) and the Boltzmann initial data takes the form (3.8), and we shall take advantage of the control by stationary solutions (the existence of which is obviously related to these cancellations) given by Proposition 4.1 to obtain a lifespan which does not depend on the initial data. Indeed, we have thanks to Propositions 3.2 and 4.1 provided that $\alpha\varepsilon \ll 1$

$$\begin{aligned}\|f_N^{(k)}(t)\|_{\varepsilon,k,\beta} &= \operatorname{supess}_{Z_k\in\mathcal{D}_\varepsilon^k}\left|f_N^{(k)}(t,Z_k)\ \exp\left(\beta H_k(Z_k)\right)\right|\\ &\leq \sup_{Z_k\in\mathcal{D}_\varepsilon^k}\left(M_{N,\beta}^{(k)}(Z_k)\exp\left(\beta H_k(Z_k)\right)\right)\|\rho^0\|_{L^\infty}\\ &\leq C^k \sup_{Z_k\in\mathcal{D}_\varepsilon^k}\left(M_\beta^{\otimes k}(V_k)\exp\left(\beta H_k(Z_k)\right)\right)\|\rho^0\|_{L^\infty}.\end{aligned}$$

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

$$\|f_N^{(k)}(t)\|_{\varepsilon,k,\beta} \leq C^k\left(\frac{\beta}{2\pi}\right)^{kd/2}\|\rho^0\|_{L^\infty}\,. \tag{4.6}$$

Similarly for the initial data for the Boltzmann hierarchy defined in (3.8), by Remark 3.5 the solution (3.11) of the evolution is bounded by

$$\|g_\alpha^{(k)}(t)\|_{0,k,\beta} \leq \left(\frac{\beta}{2\pi}\right)^{kd/2} \|\rho^0\|_{L^\infty}. \tag{4.7}$$

Moreover we shall use a truncated series expansion instead of (3.5) and (3.10). Let us fix a (small) parameter $h > 0$ and a sequence $\{n_k\}_{k\geq 1}$ of integers to be tuned later. We shall study the dynamics up to time $t := Kh$ for some large integer K, by splitting the time interval $[0, t]$ into K intervals, and controlling the number of collisions on each interval. In order to discard trajectories with a large number of collisions in the iterated Duhamel formula (3.5), we define collision trees "of controled size" by the condition that they have strictly less than n_k branch points on the interval $[t - kh, t - (k-1)h]$. 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 (3.3) with $s = 1$, during the time interval $[t - h, t]$: iterating Duhamel's formula up to time $t - h$ instead of time 0, we have

$$f_N^{(1)}(t) = \sum_{j_1=0}^{n_1-1} \alpha^{j_1-1} Q_{1,1+j_1}(h) f_N^{(j_1)}(t-h) + R_{1,n_1}(t-h,t), \tag{4.8}$$

where R_{1,n_1} accounts for at least n_1 collisions

$$R_{1,n_1}(t',t) := \sum_{p=n_1}^{N-1} \alpha^p Q_{1,p+1}(t-t') f_N^{(p+1)}(t').$$

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

$$R_{k,n}(t',t) := \sum_{p=n}^{N-k} \alpha^p Q_{k,k+p}(t-t') f_N^{(k+p)}(t').$$

The term $R_{k,n}(t',t)$ accounts for trajectories originating at k points at time t, 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 (4.8) can be broken up again by iterating the Duhamel formula on the time interval $[t-2h, t-h]$ and truncating the contributions with more than n_2 collisions: this gives

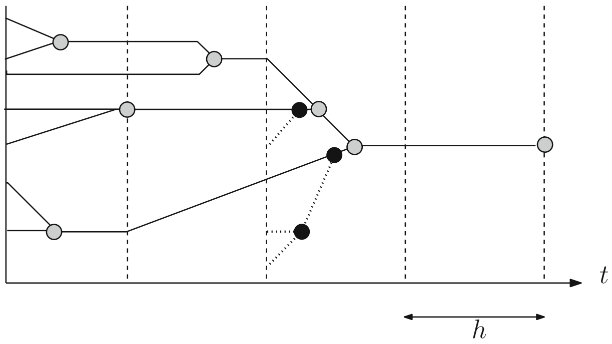

Fig. 2 Suppose $n_k = A^k$ with $A = 2$. Each collision is represented by a *circle* from which 2 trajectories emerge. The tree including the three extra collisions in *dotted lines* occurring during $[t-2h, t-h]$ is not a good collision tree and in our procedure, it would be truncated at time $t-2h$. The tree without the *dotted lines* is a good collision tree with $t = 4h$: the number of collisions during the kth-time interval is less than $n_k - 1 = A^k - 1$

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

Iterating this procedure K times and truncating the trajectories with at least n_k collisions during the time interval $[t-kh, t-(k-1)h]$, leads to the following expansion (Fig. 2)

$$f_N^{(1)}(t) = f_N^{(1,K)}(t) + R_N^K(t), \tag{4.9}$$

where denoting $J_0 := 1$ and $J_k := 1 + j_1 + \cdots + j_k$,

$$f_N^{(1,K)}(t) := \sum_{j_1=0}^{n_1-1} \cdots \sum_{j_K=0}^{n_K-1} \alpha^{J_K-1} Q_{1,J_1}(h) Q_{J_1,J_2}(h) \ldots Q_{J_{K-1},J_K}(h)\, f_N^{0(J_K)} \tag{4.10}$$

and

$$R_N^K(t) := \sum_{k=1}^{K} \sum_{j_1=0}^{n_1-1} \cdots \sum_{j_{k-1}=0}^{n_{k-1}-1} \alpha^{J_{k-1}-1} Q_{1,J_1}(h) \ldots Q_{J_{k-2},J_{k-1}}(h)$$
$$\circ R_{J_{k-1},n_k}(t-kh, t-(k-1)h).$$

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 $f_N^{(1,K)}(t)$ and that $R_N^K(t)$ vanishes asymptotically.

Next as in (4.10) we can write a truncated expansion for g_α [see (3.11)] as follows:

$$g_\alpha(t) = g_\alpha^{(1,K)}(t) + R_\alpha^{0,K}(t), \tag{4.11}$$

where with notation (3.8) and (3.11),

$$g_\alpha^{(1,K)}(t) := \sum_{j_1=0}^{n_1-1} \dots \sum_{j_K=0}^{n_K-1} \alpha^{J_K-1} Q^0_{1,J_1}(h) Q^0_{J_1,J_2}(h) \dots Q^0_{J_{K-1},J_K}(h)\, g_\alpha^{0(J_K)} \tag{4.12}$$

and

$$R_\alpha^{0,K}(t) := \sum_{k=1}^{K} \sum_{j_1=0}^{n_1-1} \dots \sum_{j_{k-1}=0}^{n_{k-1}-1} \alpha^{J_{k-1}-1} Q^0_{1,J_1}(h) \dots Q^0_{J_{k-2},J_{k-1}}(h) \circ R^0_{J_{k-1},n_k}(t-kh, t-(k-1)h)$$

with

$$R^0_{k,n}(t',t) := \sum_{p\geq n} \alpha^p Q^0_{k,k+p}(t-t') g_\alpha^{(k+p)}(t').$$

4.4 Estimates of the remainders

Since we expect the particles to undergo on average one collision per unit of time, the growth of collision trees is typically exponential. Pathological trees are therefore those with super exponential growth. There are two natural ways of defining such pathological trees

- Either by choosing some fixed h (given for instance by Lanford's proof) and $\log n_k \gg k$;
- Or by fixing $n_k = A^k$ and letting the elementary time interval $h \to 0$.

We shall choose the latter option.

Proposition 4.3 *Under the assumptions of Theorem 2.2, the following holds. Let $A \geq 2$ be given and define $n_k := A^k$, for $k \geq 1$. Then there exist $c, C, \gamma_0 > 0$ depending on d, A and β such that for any $t > 1$ and any $\gamma \leq \gamma_0$, choosing*

$$h \leq \frac{c\gamma}{\alpha^{A/(A-1)} t^{1/(A-1)}} \quad \textit{and} \quad K = t/h \textit{ integer} \tag{4.13}$$

we get

$$\left\| R_N^K(t) \right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} + \left\| R_\alpha^{0,K}(t) \right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} \leq C\gamma^A \|\rho^0\|_{L^\infty}. \quad (4.14)$$

Proof We are going to bound

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

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

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

Applying Lemma 4.2, one has (denoting generically by C_d any constant depending only on d)

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

Then arguing as in the proof of Lemma 4.2, one can write

$$\begin{aligned}&\alpha^{J_{k-1}-1}\left\| |Q|_{1,J_{k-1}}((k-1)h)\, R_{J_{k-1},n_k}(t-kh, t-(k-1)h) \right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)}\\ &\quad\leq \sum_{p=n_k}^{N-J_{k-1}} \left(\frac{C_d\alpha(k-1)h}{\beta^{(d+1)/2}}\right)^{J_{k-1}-1} \left(\frac{C_d\alpha h}{\beta^{(d+1)/2}}\right)^{p} \sup_{t\geq 0} \|f_N^{(J_{k-1}+p)}(t)\|_{\varepsilon,J_{k-1}+p,\beta}\\ &\quad\leq \|\rho^0\|_{L^\infty}\beta^{\frac{d}{2}}(\alpha t)^{J_{k-1}-1} \sum_{p=n_k}^{N-J_{k-1}} \left(\frac{C_d}{\sqrt{\beta}}\right)^{J_{k-1}+p-1} (\alpha h)^p,\end{aligned}$$

thanks to (4.6) and recalling that $(k-1)h \leq t$. Assuming from now on that

$$\frac{C_d\alpha h}{\sqrt{\beta}} < \frac{1}{2} \quad (4.15)$$

we find

$$\begin{aligned}&\alpha^{J_{k-1}-1}\left\| |Q|_{1,J_{k-1}}((k-1)h)\, R_{J_{k-1},n_k}(t-kh, t-(k-1)h) \right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)}\\ &\quad\leq \|\rho^0\|_{L^\infty}\beta^{\frac{d}{2}}(\alpha t)^{J_{k-1}-1} \left(\frac{C_d}{\sqrt{\beta}}\right)^{J_{k-1}+n_k-1} (\alpha h)^{n_k}. \quad (4.16)\end{aligned}$$

Note that $\mathcal{N}_j := 1 + n_1 + \cdots + n_j = \frac{A^{j+1}-1}{A-1} \leq \frac{1}{A-1} n_{j+1}$. Then, since $J_{k-1} \leq \mathcal{N}_{k-1}$, one has, for some appropriate constant $C(d, \beta)$,

$$\alpha^{J_{k-1}-1} \big\| |Q|_{1,J_{k-1}}((k-1)h)\, R_{J_{k-1},n_k}(t-kh, t-(k-1)h) \big\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)}$$
$$\leq \beta^{d/2} \exp\Big(A^k\Big(\log C(d,\beta) + \frac{1}{A-1}\log(\alpha t) + \log(\alpha h)\Big)\Big)\|\rho^0\|_{L^\infty}.$$

Therefore, choosing

$$h \leq \frac{\gamma}{C(d,\beta)\,\alpha^{A/(A-1)} t^{1/(A-1)}},$$

which is compatible with (4.15) as soon as γ is small enough one has

$$\alpha^{J_{k-1}-1} \big\| |Q|_{1,J_{k-1}}((k-1)h)\, R_{J_{k-1},n_k}(t-kh, t-(k-1)h) \big\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)}$$
$$\leq \beta^{d/2} \exp\Big(A^k \log\gamma\Big)\|\rho^0\|_{L^\infty}. \tag{4.17}$$

This implies

$$\begin{aligned}
\left\| R_N^K \right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} &\leq \beta^{d/2} \sum_{k=1}^K \Big(\prod_{i=1}^k n_i\Big) \exp\Big(A^k \log\gamma\Big)\|\rho^0\|_{L^\infty} \\
&\leq \beta^{d/2} \sum_{k=1}^K \exp\Big(k(k+1)\log(A) + A^k\log\gamma\Big)\|\rho^0\|_{L^\infty} \\
&\leq C_A \beta^{d/2} \sum_{k=1}^K \exp\Big(Ak\log\gamma\Big)\|\rho^0\|_{L^\infty} \leq C_A\beta^{d/2}\gamma^A\|\rho^0\|_{L^\infty}
\end{aligned}$$

for γ sufficiently small, where C_A is a constant depending on A. Thus, we get the first part of (4.14)

$$\|R_N^K\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} \leq C\gamma^A \|\rho^0\|_{L^\infty}.$$

The argument is identical in the case of the Boltzmann hierarchy:

$$\alpha^{J_{k-1}-1} \Big\| |Q|^0_{1,J_{k-1}}((k-1)h)\, R^0_{J_{k-1},n_k}(t-kh, t-(k-1)h) \Big\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)}$$
$$\leq \left(\frac{C_d(k-1)\alpha h}{\beta^{(d+1)/2}}\right)^{J_{k-1}-1} \left(\frac{C_d \alpha h}{\beta^{(d+1)/2}}\right)^{n_k} \sup_{t\geq 0} \|g_\alpha^{(J_{k-1}+n_k)}(t)\|_{0,J_{k-1}+n_k,\beta}$$
$$\leq \beta^{\frac{d}{2}} \left(\frac{C_d}{\sqrt{\beta}}\right)^{J_{k-1}+n_k-1} (\alpha t)^{J_{k-1}-1} (\alpha h)^{n_k} \|\rho^0\|_{L^\infty},$$

hence finally

$$\left\| R_\alpha^{0,K} \right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} \le C_A \beta^{d/2}\gamma^A \|\rho^0\|_{L^\infty},$$

and the proposition is proved. □

5 Proof of the convergence

In this section, we conclude the proof of Theorem 2.2. Thanks to Proposition 4.3, we are reduced to studying $f_N^{(1,K)} - g_\alpha^{(1,K)}$ [introduced in (4.9), (4.11)] and to proving that the matching terms in the series $f_N^{(1,K)}$ and $g_\alpha^{(1,K)}$ are close to each other.

Throughout this section, the parameters are chosen such that (with the notation of Proposition 4.3)

$$N\varepsilon^{d-1} = \alpha \ll \frac{1}{\varepsilon}, \quad A \ge 2, \quad t > 1, \quad K = \frac{t}{h}. \tag{5.1}$$

Each elementary term in the series $f_N^{(1,K)}$ and $g_\alpha^{(1,K)}$ has a geometric interpretation as an integral over some *pseudo-trajectories*. As explained in [13,21,28], in this formulation the characteristics associated with the operators $\mathbf{S}_i(t_{i-1}-t_i)$ and $\mathbf{S}_i^0(t_{i-1}-t_i)$ are followed backwards in time between two consecutive times t_i and t_{i-1}, and the collision terms (associated with $C_{i,i+1}$ and $C_{i,i+1}^0$) are seen as source terms in which "additional particles" are "adjoined" to the system. The main heuristic idea is that the pseudo-trajectories associated to both hierarchies can be coupled precisely if no recollisions occur in the BBGKY hierarchy. The core of the proof will be to obtain an upper bound on the occurrence of recollisions and to show that their contribution is negligible.

In order to prevent recollisions in the time interval $[t_{i+1}, t_i]$, some bad sets in phase space must be removed. Following the approach developed in [21], a geometrical control of the trajectories in the torus (stated in Lemma 5.2) enables us to define bad sets, outside of which the flow $\mathbf{S}$ between two collision times is the free flow $\mathbf{S}^0$ (see Proposition 5.1). Finally, the geometric controls are used in Sect. 5.3 to obtain quantitative estimates on the collision integrals where those bad sets have been removed.

5.1 Reformulation in terms of pseudo-trajectories

We consider one term of the sum $f_N^{(1,K)}(t)$ in (4.10) and show how it can be interpreted in terms of pseudo-trajectories. Given the indices

$J = (j_1, \dots, j_K)$, we set

$$\begin{aligned} & F_N^{(1,K)}(J)\,(t, z_1) \\ & := Q_{1,J_1}(h) Q_{J_1,J_2}(h) \dots Q_{J_{K-1},J_K}(h)\, f_N^{0(J_K)} \\ & = \int_{\mathcal{T}_J(h)} dT\; \mathbf{S}_1(t - t_1) C_{1,2} \mathbf{S}_2(t_1 - t_2) C_{2,3} \dots \mathbf{S}_{J_K}(t_{J_K - 1}) f_N^{0(J_K)} \end{aligned} \tag{5.2}$$

where the time integral is over the collision times $T = (t_1, \dots, t_{J_{K-1}})$ taking values in

$$\mathcal{T}_J(h) := \Big\{ T = (t_1, \dots, t_{J_{K-1}}) \,\Big|\, t_i < t_{i-1} \text{ and } (t_{J_k}, \dots, t_{J_{k-1}+1}) \in [t - kh, t - (k-1)h] \Big\}. \tag{5.3}$$

In the following we denote by $\mathbf{\Psi}_s$ the s-particle flow. Given $z_1 = (x_1, v_1) \in \mathbf{T}^d \times \mathbf{R}^d$ and a time $u \in [t_1, t]$, we call $z_1(u) = \mathbf{\Psi}_1(u) z_1$ the coordinates following the backward flow $\mathbf{\Psi}_1$ of one particle. The first collision operator $C_{1,2}$ is interpreted as the adjunction at time t_1 of a new particle at $x_1(t_1) + \varepsilon \nu_2$ for a deflection angle $\nu_2 \in \mathbf{S}^{d-1}$ and with a velocity $v_2 \in \mathbf{R}^d$. The new pair of particles Z_2 will be evolving according to the backward 2-particle flow $\mathbf{\Psi}_2$ during the time interval $[t_2, t_1]$ starting at t_1 from

$$\begin{cases} Z_2(t_1) = \big((x_1(t_1), v_1), (x_1(t_1) + \varepsilon \nu_2, v_2) \big) \text{ in the pre-collisional case } (v_2 - v_1) \cdot \nu_2 < 0 \\ Z_2(t_1) = \big((x_1(t_1), v_1^*), (x_1(t_1) + \varepsilon \nu_2, v_2^*) \big) \text{ in the post-collisional case } (v_2 - v_1) \cdot \nu_2 > 0, \end{cases} \tag{5.4}$$

the latter case corresponding to the scattering.

Iterating this procedure, a branching process is built inductively by adding a particle labelled $i + 1$ at time t_i to the particle $z_{m_i}(t_i)$ where $m_i \le i$ is chosen randomly among the first i particles. Given a deflection angle ν_{i+1} and a velocity v_{i+1}, the velocity of the particles z_{m_i} and z_{i+1} at time t_i are updated according to the pre-collisional or post-collisional rule as in (5.4)

$$\begin{cases} Z_{i+1}(t_i) = \big(\{z_j(t_i)\}_{j \neq m_i}, (x_{m_i}(t_i), v_{m_i}(t_i)), (x_{m_i}(t_i) + \varepsilon \nu_{i+1}, v_{i+1}) \big) \\ \qquad \text{in the pre-collisional case} (v_{i+1}(t_i) - v_{m_i}) \cdot \nu_{i+1} < 0 \\ Z_{i+1}(t_i) = \big(\{z_j(t_i)\}_{j \neq m_i}, (x_{m_i}(t_i), v^*_{m_i}(t_i)), (x_{m_i}(t_i) + \varepsilon \nu_{i+1}, v^*_{i+1}) \big) \\ \qquad \text{in the post-collisional case } (v_{i+1}(t_i) - v_{m_i}) \cdot \nu_{i+1} > 0. \end{cases}$$

Let Z_{i+1} denote the $i + 1$ components after the i^{th}-collision. The evolution of Z_{i+1} follows the flow of the backward transport $\mathbf{\Psi}_{i+1}$ during the time

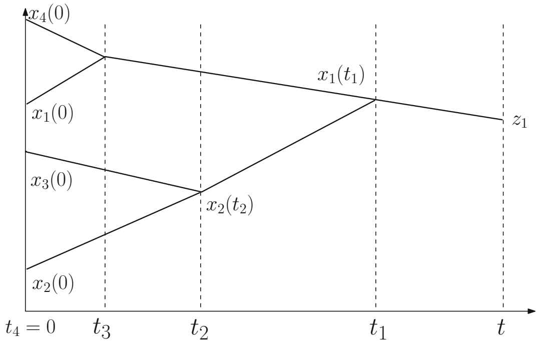

Fig. 3 A collision tree is represented with 3 collisions. The velocities (v_1, v_2) at time t_1 are pre-collisional and the first particle keeps its velocity v_1 after the collision. At time t_3, the first particle is selected $m_3 = 1$ and the velocity v_1 is modified into v_1^* according to the post-collisional rule

interval $[t_{i+1}, t_i]$. From [40] (see also Remark 3.1), one can check that $\mathbf{\Psi}_{i+1}$ is well defined up to a set of measure 0. In the following, we shall use the name *collision* to describe the creation of a particle and *recollision* if two particles collide in the flow $\mathbf{\Psi}_{i+1}$.

To summarize, pseudo-trajectories do not involve physical particles. They are a geometric interpretation of the iterated Duhamel formula in terms of a branching process flowing backward in time (Fig. 3) and determined by

- The collision times $T = (t_1, \dots, t_{J_K - 1})$ which are interpreted as branching times
- The labels of the collision particles $m = (m_1, \dots, m_{J_K - 1})$ from which branching occurs and which take values in the set

$$\mathcal{M}_J := \left\{ m = (m_1, \dots, m_{J_K - 1}), \quad 1 \le m_i \le i \right\}$$

- The coordinates of the initial particle z_1 at time t
- The velocities $v_2, \dots, v_{J_K}$ in $\mathbf{R}^d$ and deflection angles $\nu_2, \dots, \nu_{J_K}$ in $\mathbf{S}_1^{d-1}$ for each additional particle.

The integral (5.2) can be evaluated by integrating $f_N^{0(J_K)}$ on the value of the pseudo-trajectories $Z_{J_K}(0)$ at time 0

$$F_N^{(1,K)}(J) = \sum_{m \in \mathcal{M}_J} \left(\frac{\varepsilon^{d-1}}{\alpha} \right)^{J_K - 1} \frac{(N-1)!}{(N - J_K)!} \, F_N^{(1,K)}(J, m)$$

where

$$F_N^{(1,K)}(J, m)\,(t, z_1) := \int_{\mathcal{T}_J(h)} dT \int_{(\mathbf{S}^{d-1} \times \mathbf{R}^d)^{J_K - 1}} d\bar{\nu}\, d\bar{V} \\ \mathcal{A}(T, z_1, \bar{\nu}, \bar{V}) \; f_N^{0(J_K)}(Z_{J_K}(0)) \qquad (5.5)$$

and with

$$\mathcal{A}(T, z_1, \bar{\nu}, \bar{V}) := \prod_{i=1}^{J_K - 1} ((v_{i+1} - v_{m_i}(t_i)) \cdot \nu_{i+1}) \text{ and } \begin{cases} \bar{\nu} = \{\nu_2, \dots, \nu_{J_K}\} \\ \bar{V} = \{v_2, \dots, v_{J_K}\} \end{cases}. \tag{5.6}$$

The definition of $\mathcal{A}$ requires to compute the whole pseudo-trajectory on the time interval $[0, t]$ starting at z_1 in order to be able to sample the velocities at the different times $T = (t_1, \dots, t_{J_K - 1})$. Note that the contributions of the gain and loss terms in the collision operator $C_{k,k+1}$ are taken into account by the sign of $\big((v_{k+1} - v_{m_k}(t_k)) \cdot \nu_{k+1}\big)$.

In the same way, a branching process associated with the Boltzmann hierarchy can be constructed: given an initial particle $z_1^0 = (x_1^0, v_1^0)$ at time t, a collection of collision times $T = (t_1, \dots, t_{J_K - 1})$ and labels of the collision particles $m = (m_1, \dots, m_{J_K - 1}) \in \mathcal{M}_J$ as well as a collection of velocities $v_2, \dots, v_{J_K}$ and deflection angles $\nu_2, \dots, \nu_{J_K}$, the $(k+1)^{\text{th}}$ particle z_{k+1}^0 is added at time t_k at the position $x_{m_k}^0(t_k)$ of the particle m_k and their velocities are adjusted according to the type of the collision

$$\begin{cases} z_{m_k}^0(t_k) = \big(x_{m_k}^0(t_k), v_{m_k}(t_k)\big), \; z_{k+1}^0(t_k) = \big(x_{m_k}^0(t_k), v_{k+1}\big) \\ \quad \text{if } (v_{k+1} - v_{m_k}(t_k)) \cdot \nu_{k+1} < 0 \\ z_{m_k}^0(t_k) = \big(x_{m_k}^0(t_k^+), v_{m_k}^*(t_k^+)\big), \; z_{k+1}^0(t_k) = \big(x_{m_k}^0(t_k^+), v_{k+1}^*\big) \\ \quad \text{if } (v_{k+1} - v_{m_k}(t_k^+)) \cdot \nu_{k+1} > 0. \end{cases}$$

Then, the corresponding pseudo-trajectory Z_{k+1}^0 evolves according to the backward free flow denoted by $\mathbf{\Psi}_{k+1}^0$ during the time interval $[t_{k+1}, t_k]$ until the next particle creation. As the particles are points, no recollision occurs in this branching process. Notice that $u \mapsto Z_{k+1}^0(u)$ is pointwise left-continuous on $[0, t_k]$.

The counterpart of the integral (5.2) in the series $g_\alpha^{(1,K)}(t)$ in (4.12) can be formally rewritten as follows

$$\begin{aligned} G^{(1,K)}(J)\,(t, z_1) &= \int_{\mathcal{T}_J(h)} dT\; \mathbf{S}_1^0(t - t_1) C_{1,2}^0 \mathbf{S}_2^0(t_1 - t_2) C_{2,3}^0 \dots \mathbf{S}_{J_K}^0(t_{J_K - 1}) g^{0(J_K)} \\ &= \sum_{m \in \mathcal{M}_J} G^{(1,K)}(J, m) \end{aligned} \tag{5.7}$$

where the integral is over the pseudo-trajectories

$$\begin{aligned} &G^{(1,K)}(J, m)\,(t, z_1) \\ &\quad := \int_{\mathcal{T}_J(h)} dT \int_{(\mathbf{S}^{d-1} \times \mathbf{R}^d)^{J_K - 1}} d\bar{\nu}\, d\bar{V}\; \hat{\mathcal{A}}(T, z_1, \bar{\nu}, \bar{V})\; g^{0(J_K)}(Z_{J_K}^0(0)), \end{aligned} \tag{5.8}$$

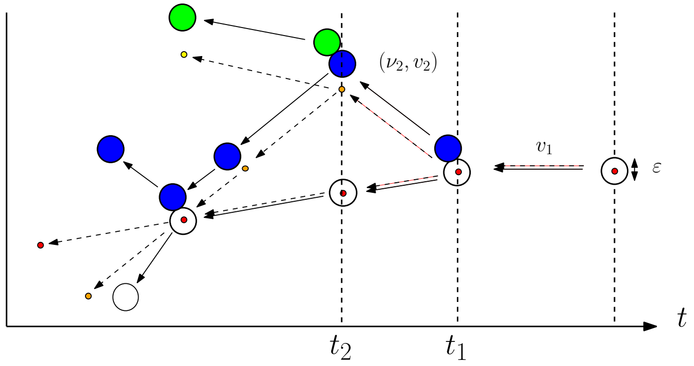

Fig. 4 The first stages of both pseudo-trajectories are depicted up to the occurence of a recollision. The BBGKY pseudo-trajectories are represented with *plain arrows*, whereas the Boltzmann pseudo-trajectories correspond to the *dashed arrows*. At time t, the particle with label 1 in the BBGKY hierarchy is a ball of radius ε centered at position x_1 and the particle in the Boltzmann hierarchy is depicted as a point located at $x_1^0 = x_1$. At time t_1 the second particle is added and at time t_2 the third. Both hierarchies are coupled, but a small error in the particle positions of order ε can occur at each collision. In this figure, a recollision between the first and the second particle of the BBGKY pseudo-trajectories occurs and after this recollision the Boltzmann and the BBGKY pseudo-trajectories are no longer close to each other. Indeed the BBGKY trajectories are deflected after the recollision, instead the ideal particles do not collide and follow a straight line (see the *dashed arrows*). Note that before the recollisions the trajectories of z_1 and z_1^0 are identical and therefore the plain and the *dashed arrows* overlap

with $\hat{\mathcal{A}}$ defined as in (5.6) but with respect to the Boltzmann hierarchy pseudo-trajectories.

To show that $F_N^{(1,K)}(J, m)$ and $G^{(1,K)}(J, m)$ are close to each other when N diverges, we shall prove that the pseudo-trajectories Z and Z^0 can be coupled in order to remain very close to each other up to a small error (see Fig. 4)

- Due to the micro-translations $\varepsilon \nu_{k+1}$ of the added particle at each collision time t_k
- Excluding the possible recollisions on the interval $]t_k, t_{k-1}[$ along the flow $\mathbf{S}_k$, which do not occur for the free flow $\mathbf{S}_k^0$.

The proof of the convergence follows the arguments of [21]. This will be achieved by constructing in (5.18), a set of deflection angles and velocities $(\mathcal{B}(J, T, m))^c \subset \left(\mathbf{S}^{d-1} \times \mathbf{R}^d\right)^{J_k - 1}$ such that the pseudo-trajectories Z induced by this set have no recollisions and therefore remain very close to the pseudo-trajectories Z^0 associated to the free flow. Furthermore, the measure of $\mathcal{B}(J, T, m)$ tends to 0 when N goes to infinity. Finally, in Sect. 5.3, all the estimates will be combined to derive a quantitative bound on $F_N^{(1,K)}(J, m) - G^{(1,K)}(J, m)$.

5.2 Reduction to non-pathological trajectories

5.2.1 The elementary step

The set of good configurations with k particles will be such that the particles remain at a distance $\varepsilon_0 \gg \varepsilon$ for a time t, i.e. that they belong to the set

$$\mathcal{G}_k(\varepsilon_0) := \Big\{ Z_k \in \mathbf{T}^{dk} \times \mathbf{R}^{dk} \;\Big|\; \forall u \in [0,t], \quad \forall i \neq j, \\ d(x_i - u\, v_i, x_j - u\, v_j) \geq \varepsilon_0 \Big\}$$

where d denotes the distance on the torus $\mathbf{T}^d$. For particles in $\mathcal{G}_k(\varepsilon_0)$, the transport $\mathbf{\Psi}_k$ coincides with the free flow. Fix $\bar{a} \ll \varepsilon_0$. Thus, if at time t the configurations Z_k, Z_k^0 are such that

$$\forall i \leq k, \qquad |x_i - x_i^0| \leq \bar{a}, \qquad v_i = v_i^0 \tag{5.9}$$

and that Z_k^0 belongs to $\mathcal{G}_k(\varepsilon_0)$, then the configurations $\mathbf{\Psi}_k(u) Z_k$, $\mathbf{\Psi}_k^0(u) Z_k^0$ will remain at distance less than $\bar{a}$ for $u \in [0,t]$.

We are going to show that the good configurations are stable by adjunction of a $(k+1)^{\text{th}}$-particle next to the particle labelled by $m_k \leq k$. More precisely, let $Z_k^0 = (X_k^0, V_k)$ be in $\mathcal{G}_k(\varepsilon_0)$ and $Z_k = (X_k, V_k)$ with positions close to X_k^0 and same velocities [cf. (5.9)]. Then, by choosing the velocity v_{k+1} and the deflection angle ν_{k+1} of the new particle $k+1$ outside a bad set $\mathcal{B}_k^{m_k}(Z_k^0)$, both configurations Z_k and Z_k^0 will remain close to each other. Of course, immediately after the adjunction, the particles m_k and $k+1$ will not be at distance ε_0, but v_{k+1}, ν_{k+1} will be chosen such that the particles drift rapidly far apart and after a short time $\delta > 0$ the configurations Z_{k+1} and Z_{k+1}^0 will be again in the good sets $\mathcal{G}_{k+1}(\varepsilon_0/2)$ and $\mathcal{G}_{k+1}(\varepsilon_0)$.

This stability result was obtained in [21] and is stated below. We shall restrict to bounded velocities taking values in the ball $B_E := \big\{ v \in \mathbf{R}^d,\ |v| \leq E \big\}$ for a given large parameter $E > 0$ to be tuned later on.

Proposition 5.1 ([21]) *We fix parameters $\bar{a}, \varepsilon_0, \delta$ such that*

$$A^{K+1}\varepsilon \ll \bar{a} \ll \varepsilon_0 \ll \min(\delta E, 1). \tag{5.10}$$

Given $Z_k^0 = (X_k^0, V_k) \in \mathcal{G}_k(\varepsilon_0)$ and $m_k \leq k$, there is a subset $\mathcal{B}_k^{m_k}(Z_k^0)$ of $\mathbf{S}^{d-1} \times B_E$ of small measure

$$\big|\mathcal{B}_k^{m_k}(Z_k^0)\big| \leq Ck \left(E^d \left(\frac{\bar{a}}{\varepsilon_0}\right)^{d-1} + E^d (Et)^d \varepsilon_0^{d-1} + E \left(\frac{\varepsilon_0}{\delta}\right)^{d-1} \right) \tag{5.11}$$

such that good configurations close to Z_k^0 are stable by adjunction of a collisional particle close to the particle $x_{m_k}^0$ in the following sense.

Let $Z_k = (X_k, V_k)$ be a configuration of k particles satisfying (5.9), i.e. $|X_k - X_k^0| \leq \bar{a}$. Given $(\nu_{k+1}, v_{k+1}) \in (\mathbf{S}^{d-1} \times B_E) \setminus \mathcal{B}_k^{m_k}(Z_k^0)$, a new particle with velocity v_{k+1} is added at $x_{m_k} + \varepsilon \nu_{k+1}$ to Z_k and at $x_{m_k}^0$ to Z_k^0. Two possibilities may arise

- *For a pre-collisional configuration $\nu_{k+1} \cdot (v_{k+1} - v_{m_k}) < 0$ then*

$$\forall u \in]0, t], \quad \begin{cases} \forall i \neq j \in [1, k], \quad d(x_i - u\, v_i, x_j - u\, v_j) > \varepsilon, \\ \forall j \in [1, k], \quad d(x_{m_k} + \varepsilon \nu_{k+1} - u\, v_{k+1}, x_j - u\, v_j) > \varepsilon. \end{cases} \tag{5.12}$$

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

$$\forall u \in [\delta, t], \quad \begin{cases} (X_k - u V_k, V_k, x_{m_k} + \varepsilon \nu_{k+1} - u\, v_{k+1}, v_{k+1}) \in \mathcal{G}_{k+1}(\varepsilon_0/2) \\ (X_k^0 - u V_k, V_k, x_{m_k}^0 - u\, v_{k+1}, v_{k+1}) \in \mathcal{G}_{k+1}(\varepsilon_0). \end{cases} \tag{5.13}$$

- *For a post-collisional configuration $\nu_{k+1} \cdot (v_{k+1} - v_{m_k}) > 0$ then the velocities are updated*

$$\forall u \in]0, t], \begin{cases} \forall i \neq j \in [1, k] \setminus \{m_k\}, \quad d(x_i - u\, v_i, x_j - u\, v_j) > \varepsilon, \\ \forall j \in [1, k] \setminus \{m_k\}, \quad d(x_{m_k} + \varepsilon \nu_{k+1} - u\, v_{k+1}^*, x_j - u\, v_j) > \varepsilon, \\ \forall j \in [1, k] \setminus \{m_k\}, \quad d(x_{m_k} - u\, v_{m_k}^*, x_j - u\, v_j) > \varepsilon, \\ d(x_{m_k} - u\, v_{m_k}^*, x_{m_k} + \varepsilon \nu_{k+1} - u\, v_{k+1}^*) > \varepsilon. \end{cases} \tag{5.14}$$

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

$$\forall u \in [\delta, t], \quad \begin{cases} \big(\{x_j - u\, v_j, v_j\}_{j \neq m_k}, x_{m_k} - u\, v_{m_k}^*, v_{m_k}^*, x_{m_k} + \varepsilon \nu_{k+1} - u\, v_{k+1}^*, v_{k+1}^* \big) \\ \qquad \in \mathcal{G}_{k+1}(\varepsilon_0/2), \\ \big(\{x_j^0 - u\, v_j, v_j\}_{j \neq m_k}, x_{m_k}^0 - u\, v_{m_k}^*, v_{m_k}^*, x_{m_k}^0 - u\, v_{k+1}^*, v_{k+1}^* \big) \in \mathcal{G}_{k+1}(\varepsilon_0). \end{cases} \tag{5.15}$$

Proposition 5.1 is the elementary step for adding a new particle. In Sect. 5.2.2, we are going to show how this step can be iterated in order to build inductively good pseudo-trajectories Z and Z^0. Note that after adding a new particle, the velocities remain identical at each time in both configurations, but their positions differ due the exclusion condition in the BBGKY hierarchy which induces a shift of ε at each creation of a new particle (see Fig. 4).

We refer to [21] for a complete proof of Proposition 5.1 and simply recall that it can be obtained from the following control on free trajectories.

Lemma 5.2 *Given $t > 0$, and $\bar{a} > 0$ satisfying $A^{K+1}\varepsilon \ll \bar{a} \ll \varepsilon_0 \ll \min(\delta E, 1)$, consider two points x_1^0, x_2^0 in $\mathbf{T}^d$ such that $d(x_1^0, x_2^0) \geq \varepsilon_0$, and a velocity $v_1 \in B_E$. Then there exists a subset $K(x_1^0 - x_2^0, \varepsilon_0, \bar{a})$ of $\mathbf{R}^d$ with measure bounded by*

$$|K(x_1^0 - x_2^0, \varepsilon_0, \bar{a})| \leq C E^d \left(\left(\frac{\bar{a}}{\varepsilon_0} \right)^{d-1} + (Et)^d \, \bar{a}^{d-1} \right)$$

and a subset $K_\delta(x_1^0 - x_2^0, \varepsilon_0, \bar{a})$ of $\mathbf{R}^d$, the measure of which satisfies

$$|K_\delta(x_1^0 - x_2^0, \varepsilon_0, \bar{a})| \leq C E \left(\left(\frac{\varepsilon_0}{\delta} \right)^{d-1} + (Et)^d E^{d-1} \varepsilon_0^{d-1} \right)$$

such that for any $v_2 \in B_E$ and x_1, x_2 such that $|x_1 - x_1^0| \leq \bar{a}$, $|x_2 - x_2^0| \leq \bar{a}$, the following results hold:

- *If $v_1 - v_2 \notin K(x_1^0 - x_2^0, \varepsilon_0, \bar{a})$, then*

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

- *If $v_1 - v_2 \notin K_\delta(x_1^0 - x_2^0, \varepsilon_0, \bar{a})$*

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

The proof of this lemma is a simple adaptation of Lemma 12.2.1 in [21], 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 $\mathbf{T}^d$ 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.

5.2.2 Induction procedure for the pseudo-trajectories

Using the elementary step of Sect. 5.2.1, we are going to construct in Proposition 5.3 a coupling between the BBGKY and Boltzmann pseudo-trajectories, defined in Sect. 5.1, such that both trajectories remain close for all times up to a small error. In particular, this proof shows that recollisions may occur for the

BBGKY pseudo-trajectories only for a set of configurations at time 0 in $\mathcal{D}_\varepsilon^{J_K}$ with small measure.

As the stability of the good configurations (proved in Proposition 5.1) requires a delay $\delta > 0$ in between 2 collisions, we introduce a modified set of collision times

$$\mathcal{T}_{J,\delta}(h) := \Big\{ T = (t_1, \dots, t_{J_K-1}) \,/\, t_i < t_{i-1} - \delta,\\ (t_{J_k}, \dots, t_{J_{k-1}+1}) \in [t - kh, t - (k-1)h] \Big\}. \tag{5.16}$$

The following statement is analogous to Lemma 14.1.1 of [21].

Proposition 5.3 *Fix $J = (j_1, \dots, j_K)$, $m = (m_1, \dots, m_{J_K-1}) \in \mathcal{M}_J$ and $T \in \mathcal{T}_{J,\delta}(h)$. Let the pseudo-trajectories $Z_i = (X_i, V_i)$, $Z_i^0 = (X_i^0, V_i)$ be defined inductively by choosing at each collision time t_i a deflection angle ν_{i+1} and a velocity v_{i+1} such that*

$$(\nu_{i+1}, v_{i+1}) \in \big(\mathbf{S}^{d-1} \times B_E\big) \setminus \mathcal{B}_i^{m_i}(Z_i^0(t_i)) \quad \text{and} \quad \sum_{k=1}^{i+1} v_k^2 < E^2.$$

The velocities of both pseudo-trajectories coincide as well as the positions $x_1(u) = x_1^0(u)$ for $u \in [0, t]$. Furthermore, for ε sufficiently small

$$\forall i \le J_K - 1,\ \forall \ell \le i+1, \qquad |x_\ell(t_{i+1}) - x_\ell^0(t_{i+1})| \le \varepsilon i. \tag{5.17}$$

As a consequence of this proposition, we define a bad set of velocities and deflection angles for the pathological pseudo-trajectories

$$\mathcal{B}(z_1, J, T, m) := \Bigg\{ (\nu_i, v_i)_{2 \le i \le J_K} \in \big(\mathbf{S}^{d-1} \times B_E\big)^{J_K - 1} \,\Big|\, \sum_{k=1}^{J_K} v_k^2 < E^2 \\ \text{and} \quad \exists i_0 \le J_K - 1 \\ \text{such that} \quad \forall i < i_0, \quad (\nu_{i+1}, v_{i+1}) \in \big(\mathcal{B}_i^{m_i}(Z_i^0(t_i))\big)^c \\ \text{and} \quad (\nu_{i_0+1}, v_{i_0+1}) \in \mathcal{B}_{i_0}^{m_{i_0}}(Z_{i_0}^0(t_{i_0})) \Bigg\}. \tag{5.18}$$

Proof We proceed by induction on i, the index of the time variables t_i for $1 \le i \le J_K - 1$. The recursion hypothesis at step i is

$$Z_i^0(t_i) \in \mathcal{G}_i(\varepsilon_0) \quad \text{and} \quad \forall \ell \le i, \quad |x_\ell(t_i) - x_\ell^0(t_i)| \le \varepsilon(i-1), \quad v_\ell(t_i) = v_\ell^0(t_i). \tag{5.19}$$

We first notice that by construction, $z_1(t_1) = z_1^0(t_1)$, so (5.19) holds for $i = 1$. The initial configuration containing only one particle, there is no possible recollision!

Assume that (5.19) holds up to some $i \leq J_K - 1$ and let us prove that (5.19) holds for $i + 1$. 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 $(v_{i+1}, v_{m_i}(t_i))$ at time t_i. We recall that the particle is adjoined in such a way that (ν_{i+1}, v_{i+1}) belongs to $\big(\mathbf{S}^{d-1} \times B_E\big) \backslash \mathcal{B}_i^{m_i}(Z_i^0(t_i))$. The new configuration Z_{i+1}^0 satisfies for all $u \in]t_{i+1}, t_i]$

$$
\begin{aligned}
\forall \ell \leq i, \quad x_\ell^0(u) &= x_\ell^0(t_i) + (u - t_i) v_\ell(t_i), & v_\ell(u) &= v_\ell(t_i), \\
x_{i+1}^0(u) &= x_{m_i}^0(t_i) + (u - t_i) v_{i+1}, & v_{i+1}(u) &= v_{i+1}.
\end{aligned}
$$

Since $t_i - t_{i+1} > \delta$, Proposition 5.1 implies that $Z_{i+1}^0(t_{i+1})$ will be in $\mathcal{G}_{i+1}(\varepsilon_0)$.

Now let us study Z_{i+1} the BBGKY pseudo-trajectory. Provided that ε is sufficiently small, by the induction assumption (5.19) and the fact that $A^{K+1}\varepsilon \leq \bar{a}$ [see (5.10)], 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 $\mathcal{G}_i(\varepsilon_0)$, Proposition 5.1 implies that backwards in time, there is free flow for Z_{i+1}. In particular,

$$
\begin{aligned}
\forall \ell < i+1, \quad x_\ell(u) &= x_\ell(t_i) + (u - t_i) v_\ell(t_i), & v_\ell(u) &= v_\ell(t_i), \\
x_{i+1}(u) &= x_{m_i}(t_i) + \varepsilon \nu_{i+1} + (u - t_i) v_{i+1}, & v_{i+1}(u) &= v_{i+1}.
\end{aligned}
\tag{5.20}
$$

Therefore, the velocities of both configurations coincide and by the induction assumption (5.19)

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

where we used that in (5.20) there is a shift by at most ε.

- The case of post-collisional velocities $(v_{i+1}, v_{m_i}(t_i))$ at time t_i is identical up to a scattering of the velocities v_{i+1}, v_{m_i} in $v_{i+1}^*, v_{m_i}^*$. Note that the constraint $\sum_{k=1}^{i+1} |v_k^2| < \frac{E^2}{2}$ implies that both velocities $v_{i+1}^*, v_{m_i}^*$ remain in B_E. This concludes the proof of Proposition 5.3. □

5.3 Estimate of the error term

We turn now to the main goal of this section and use the coupling of Proposition 5.3 between the hierachies to show that for $K \ll \log\log N$ and $\alpha t \ll (\log\log N)^{(A-1)/A}$ then

$$\| f_N^{(1,K)} - g_\alpha^{(1,K)} \|_{L^\infty([0,t]\times \mathbf{T}^d \times \mathbf{R}^d)} \to 0 \tag{5.21}$$

with an explicit rate of convergence when N diverges. The coupling of Proposition 5.3 can be implemented only for a reduced set of velocities taking values in B_E and for collision times separated at least by δ. Thus the first step will be to estimate the cost of cutting-off the large velocities and the collision time separation in (5.5) and (5.8). Then in Sect. 5.3.4, the parameters δ, E and K will be tuned and the error term evaluated.

5.3.1 Energy truncation

Given $E > 0$, define the large velocity cut-off for $f_N^{(1,K)}$ introduced in (4.10) as

$$f_{N,E}^{(1,K)} := \sum_J \varepsilon^{(d-1)(J_K-1)} \frac{(N-1)!}{(N-J_K)!} \sum_{m\in\mathcal{M}_J} F_{N,E}^{(1,K)}(J,m)$$

where $\sum_J$ stands for $\sum_{j_1=0}^{n_1-1} \dots \sum_{j_K=0}^{n_K-1}$ and the velocities in the integral (5.5) are truncated

$$F_{N,E}^{(1,K)}(J,m)\,(t,z_1) := \int_{\mathcal{T}_J(h)} dT \int_{(\mathbf{S}^{d-1}\times B_E)^{J_K-1}} d\bar{\nu}\, d\bar{V}\, \mathcal{A}(T,z_1,\bar{\nu},\bar{V}) \\ \times \mathbf{1}_{\{H_{J_K}(Z_{J_K}(0))\le \frac{E^2}{2}\}} F_N^{0(J_K)}\big(Z_{J_K}(0)\big) \tag{5.22}$$

where $\mathcal{A}$ was defined in (5.6) and $H_k(Z_k) = \frac{1}{2}\sum_{i=1}^k |v_i|^2$.

In the same way, for $g_\alpha^{(1,K)}$ in (4.12), the large velocity cut-off is defined as

$$g_{\alpha,E}^{(1,K)} := \sum_J \alpha^{J_K-1} \sum_{m\in\mathcal{M}_J} G_E^{(1,K)}(J,m)$$

where the velocities in the integral (5.8) are truncated

$$G_E^{(1,K)}(J,m)\,(t,z_1) := \int_{\mathcal{T}_J(h)} dT \int_{(\mathbf{S}^{d-1}\times B_E)^{J_K-1}} d\bar{\nu}\, d\bar{V}\, \hat{\mathcal{A}}(T,z_1,\bar{\nu},\bar{V}) \\ \times \mathbf{1}_{\{H_{J_K}(Z^0_{J_K}(0))\le \frac{E^2}{2}\}} g^{0(J_K)}(Z^0_{J_K}(0)). \tag{5.23}$$

Then, we have the following error estimate.

Proposition 5.4 *There is a constant C depending only on β and d such that, as N goes to infinity in the scaling $N\varepsilon^{d-1}\alpha^{-1} \equiv 1$, the following bounds hold:*

$$\begin{aligned}&\|f_N^{(1,K)} - f_{N,E}^{(1,K)}\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} + \|g_\alpha^{(1,K)} - g_{\alpha,E}^{(1,K)}\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)}\\ &\quad\leq A^{K(K+1)}(C\alpha t)^{A^{K+1}} e^{-\frac{\beta}{4}E^2}\|\rho^0\|_{L^\infty},\end{aligned}$$

with A, K as in Proposition 4.3.

Proof We first consider the BBGKY hierarchy. Since the kinetic energy is preserved by the transport $\mathbf{S}_k$, the difference $(f_N^{(1,K)} - f_{N,E}^{(1,K)})$ can be bounded from above by estimating the contribution of the pseudo-trajectories such that $\{H_{J_K}(Z_{J_K}(0)) \geq \frac{E^2}{2}\}$ at time 0. Note that from (4.6)

$$\begin{aligned}\|\mathbf{1}_{\{H_{J_K}(Z_{J_K})\geq\frac{E^2}{2}\}} f_N^{0(J_K)}\|_{\varepsilon,J_K,\beta/2} &\leq \|f_N^{0(J_K)}\|_{\varepsilon,J_K,\beta}\, e^{-\frac{\beta}{4}E^2}\\ &\leq C^{J_K}\, e^{-\frac{\beta}{4}E^2}\|\rho^0\|_{L^\infty}. \qquad (5.24)\end{aligned}$$

By Lemma 4.2, we get

$$\begin{aligned}&\|F_N^{(1,K)}(J,m) - F_{N,E}^{(1,K)}(J,m)\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)}\\ &\quad\leq \Big\| |Q|_{1,J_K}(t)\mathbf{1}_{\{H_{J_K}(Z_{J_K})\geq\frac{E^2}{2}\}} f_N^{0(J_K)}\Big\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)}\\ &\quad\leq \left(\frac{Ct}{(\beta/2)^{(d+1)/2}}\right)^{J_K-1} \|\mathbf{1}_{\{H_{J_K}(Z_{J_K})\geq\frac{E^2}{2}\}} f_N^{0(J_K)}\|_{\varepsilon,J_K,\beta/2}.\end{aligned}$$

It follows that

$$\|F_N^{(1,K)}(J,m) - F_{N,E}^{(1,K)}(J,m)\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \leq (Ct)^{A^{K+1}} e^{-\frac{\beta}{4}E^2}\|\rho^0\|_{L^\infty}$$

thanks to (5.24) and to the fact that $J_K \leq \mathcal{N}_K \leq A^{K+1}$. A similar estimate holds for the Boltzmann hierarchy. Summing over all possible choices of j_k proves the proposition, recalling that in the Boltzmann-Grad scaling

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

Proposition 5.4 is proved. □

5.3.2 Time separation

We choose a small parameter $\delta > 0$ such that $A^K\delta \ll h$ and estimate the error for separating the collision times by at least δ. The time cut-off of the pseudo-trajectories is defined as

$$f^{(1,K)}_{N,E,\delta} := \sum_J \varepsilon^{(d-1)(J_K-1)} \frac{(N-1)!}{(N-J_K)!} \sum_{m\in\mathcal{M}_J} F^{(1,K)}_{N,E,\delta}(J,m) \tag{5.25}$$

where the time integrals are restricted to the set $\mathcal{T}_{J,\delta}(h)$ defined in (5.16)

$$F^{(1,K)}_{N,E,\delta}(J,m)\,(t,z_1) := \int_{\mathcal{T}_{J,\delta}(h)} dT \int_{(\mathbf{S}^{d-1}\times B_E)^{J_K-1}} d\bar\nu\, d\bar V\ \mathcal{A}(T,z_1,\bar\nu,\bar V) \\ \times \mathbf{1}_{\{H_{J_K}(Z_{J_K}(0))\le\frac{E^2}{2}\}} f^{0(J_K)}_N(Z_{J_K}(0))$$

with $\mathcal{A}(t,z_1,\bar\nu,\bar V)$ as in (5.6). In the same way, for the Boltzmann hierarchy, we set

$$g^{(1,K)}_{\alpha,E,\delta} := \sum_J \alpha^{J_K-1} \sum_{m\in\mathcal{M}_J} G^{(1,K)}_{E,\delta}(J,m)$$

where the separation time cut-off is defined as

$$G^{(1,K)}_{E,\delta}(J,m)\,(t,z_1) := \int_{\mathcal{T}_{J,\delta}(h)} dT \int_{(\mathbf{S}^{d-1}\times B_E)^{J_K-1}} d\bar\nu\, d\bar V\ \hat{\mathcal{A}}(T,z_1,\bar\nu,\bar V) \\ \times \mathbf{1}_{\{H_{J_K}(Z^0_{J_K}(0))\le\frac{E^2}{2}\}} g^{0(J_K)}(Z^0_{J_K}(0)).$$

Then the following holds.

Proposition 5.5 *There is a constant C depending only on β and d such that, as N goes to infinity in the scaling $N\varepsilon^{d-1}\alpha^{-1}\equiv 1$, the following holds*

$$\|f^{(1,K)}_{N,E} - f^{(1,K)}_{N,E,\delta}\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} + \|g^{(1,K)}_{\alpha,E} - g^{(1,K)}_{\alpha,E,\delta}\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \\ \le A^{(K+2)(K+1)}(C\alpha t)^{A^{K+1}} \frac{\delta}{t}\|\rho^0\|_{L^\infty}, \tag{5.26}$$

with A, K as in Proposition 4.3.

Proof Given J,m the difference $(F^{(1,K)}_{N,E} - F^{(1,K)}_{N,E,\delta})(J,m)$ involves the integration over two consecutive times such that $|t_{i+1}-t_i|\le\delta$. This leads to a contribution $\delta t^{J_K-2}/(J_K-2)!$ instead of $t^{J_K-1}/(J_K-1)!$ and there are J_K-1 possible choices for the collision with a short time separation. Modifying accordingly the estimates of Lemma 4.2, we get for a given J

$$\Big\| \sum_{m\in\mathcal{M}_J} \big(F^{(1,K)}_{N,E} - F^{(1,K)}_{N,E,\delta}\big)(J,m)\Big\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \leq (C\alpha t)^{A^{K+1}} \frac{(A^{K+1})^2\,\delta}{t}\|\rho^0\|_{L^\infty},$$

where we used that $J_K \leq A^{K+1}$. Summing over all possible choices of j_k leads to an extra factor $A^{K(K+1)}$ as in (5.26).

A similar estimate holds in the Boltzmann case and completes the proof. □

5.3.3 Neglecting the pathological pseudo-trajectories

We now reduce the domain of integration of the velocities and deflection angles outside the set $\mathcal{B}(J,T,m)$ defined in (5.18) in order to remove the pathological pseudo-trajectories. We set

$$\tilde f^{(1,K)}_{N,E,\delta} = \sum_J \alpha^{J_K-1} \sum_{m\in\mathcal{M}_J} \left(\frac{\varepsilon^{d-1}}{\alpha}\right)^{J_K-1} \frac{(N-1)!}{(N-J_K)!}\, \tilde F^{(1,K)}_{N,E,\delta}(J,m) \tag{5.27}$$

where

$$\tilde F^{(1,K)}_{N,E,\delta}(J,m)\,(t,z_1) := \int_{\mathcal{T}_{J,\delta}(h)} dT \int_{\mathcal{B}(J,T,m)^c} d\bar v\, d\bar V\, \mathcal{A}(T,z_1,\bar v,\bar V) \times \mathbf{1}_{\{H_{J_K}(Z_{J_K}(0))\leq \frac{E^2}{2}\}} f_N^{0(J_K)}(Z_{J_K}(0)) \tag{5.28}$$

with $\mathcal{A}(t,z_1,\bar v,\bar V)$ as in (5.6). In the same way, we define

$$\tilde g^{(1,K)}_{\alpha,E,\delta} = \sum_J \alpha^{J_K-1} \sum_{m\in\mathcal{M}_J} \tilde G^{(1,K)}_{E,\delta}(J,m) \tag{5.29}$$

where the domain of integration is restricted to the complement of $\mathcal{B}(J,T,m)$

$$\tilde G^{(1,K)}_{E,\delta}(J,m)\,(t,z_1) := \int_{\mathcal{T}_{J,\delta}(h)} dT \int_{\mathcal{B}(J,T,m)^c} d\bar v\, d\bar V\, \hat{\mathcal{A}}(T,z_1,\bar v,\bar V) \times \mathbf{1}_{\{H_{J_K}(Z^0_{J_K}(0))\leq \frac{E^2}{2}\}} g^{0(J_K)}(Z^0_{J_K}(0)). \tag{5.30}$$

As a consequence of Proposition 5.1 and of the continuity estimates in Lemma 4.2, the error induced by neglecting the pathological pseudo-trajectories can be estimated from above.

Proposition 5.6 *Let* $\bar{a}$, ε_0, δ *satisfying* (5.10)*. There is a constant C depending only on* β *and d such that, as N goes to infinity in the scaling* $N\varepsilon^{d-1}\alpha^{-1} \equiv 1$*, the following holds*

$$\begin{aligned}
&\left\| g^{(1,K)}_{\alpha,E,\delta} - \widetilde{g}^{(1,K)}_{\alpha,E,\delta} \right\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} + \left\| f^{(1,K)}_{N,E,\delta} - \widetilde{f}^{(1,K)}_{N,E,\delta} \right\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \\
&\leq A^{(K+2)(K+1)} (C\alpha t)^{A^{K+1}} \left(E^d \left(\frac{\bar{a}}{\varepsilon_0}\right)^{d-1} \right. \\
&\qquad \left. + E^d (Et)^d \varepsilon_0^{d-1} + E \left(\frac{\varepsilon_0}{\delta}\right)^{d-1} \right) \|\rho^0\|_{L^\infty}.
\end{aligned}$$

Proof The proof follows the same lines as the proofs of Propositions 5.4 and 5.5. 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

$$Ck \left(E^d \left(\frac{\bar{a}}{\varepsilon_0}\right)^{d-1} + E^d \Big(Et\Big)^d \varepsilon_0^{d-1} + E \left(\frac{\varepsilon_0}{\delta}\right)^{d-1} \right) \tag{5.31}$$

according to Proposition 5.1.

For a given J, there are $J_K - 1 \leq A^{K+1}$ possible choices of the integral to be modified. Therefore, the estimate on the collision operator leads to

$$\begin{aligned}
&\Big\| \sum_{m\in\mathcal{M}_J} \big(\widetilde{F}^{(1,K)}_{N,E,\delta} - F^{(1,K)}_{N,E,\delta}\big)(J,m) \Big\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \\
&\leq (Ct)^{A^{K+1}} A^{2(K+1)} \left(E^d \left(\frac{\bar{a}}{\varepsilon_0}\right)^{d-1} + E^d (Et)^d \varepsilon_0^{d-1} + E \left(\frac{\varepsilon_0}{\delta}\right)^{d-1} \right) \|\rho^0\|_{L^\infty}
\end{aligned}$$

where as previously C depends only on d and β. The term $A^{2(K+1)}$ comes from the A^{K+1} possible choices of the integral to be modified and from the additional factor $k \leq A^{K+1}$ in (5.31).

Finally summing over all the possible choices of $J = (j_1, \ldots, j_K)$ provides the additional factor $A^{K(K+1)}$ in the estimate. Similar bounds hold also for the Boltzmann hierarchy. This completes the Proposition. □

Once the pathological pseudo-trajectories have been removed, the integrals (5.28) and (5.30) differ only by the small error on the positions $Z_{J_K}(0)$, $Z^0_{J_K}(0)$ and by the initial data $f_N^{0(J_K)}$ and $g^{0(J_K)}$. Thus, one gets

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

$$\left\| \tilde f^{(1,K)}_{N,E,\delta} - \tilde g^{(1,K)}_{\alpha,E,\delta} \right\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \leq A^{K(K+1)} (C\alpha t)^{A^{K+1}} \left(\frac{A^{2(K+1)}}{N} + \alpha\varepsilon \right) \|\rho^0\|_{L^\infty}.$$

Proof There are 2 sources of discrepancies between (5.27) and (5.29).

- *The prefactors in the collision operators* : In (5.27), the elementary collision operators have prefactors of the type $(N-k)\varepsilon^{d-1}/\alpha$ that can be replaced in the limit by 1. For fixed J_K, the corresponding error is

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

  which, combined with the bound on the collision operators, leads to an error of the form

$$A^{K(K+1)} (C\alpha t)^{A^{K+1}} \frac{A^{2(K+1)}}{N} \|\rho^0\|_{L^\infty}.$$

- *Discrepancy between $f_N^{0(J_K)}(Z_{J_K}(0))$ and $g^{0(J_K)}(Z^0_{J_K}(0))$* : First of all, we note that for the coupled pseudo-trajectories

$$g^{0(J_K)}(Z_{J_K}(0)) = g^{0(J_K)}(Z^0_{J_K}(0)).$$

Indeed, by construction both pseudo-trajectories have the same velocities and $x_1 = x_1^0$. The differences between the two configurations are only on the positions of the particles added and $g^{0(J_K)}$ is independent of these positions.

By Proposition 5.3, the initial data satisfies $Z_{J_K}(0) \in \mathcal{G}_{J_K}(\varepsilon_0/2)$. According to Proposition 3.3, we have

$$\left\| \mathbf{1}_{\mathcal{G}_{J_K}(\varepsilon_0/2)} \left(f_N^{0(J_K)} - g^{0(J_K)} \right) \right\|_{0,J_K,\beta} \leq \|\rho^0\|_{L^\infty} C^{J_K} \alpha\varepsilon.$$

Using the continuity estimate in Lemma 4.2, we then deduce that the error due to the initial data can be controlled by

$$\|\rho^0\|_{L^\infty} A^{K(K+1)} (C\alpha t)^{A^{K+1}} \alpha\varepsilon.$$

This concludes Proposition 5.7. □

5.3.4 Estimate of the main term

Finally combining the previous estimates, we get

Proposition 5.8 *For parameters satisfying* (5.1) *and such that*

$$\alpha t \ll \left(\log\log N\right)^{\frac{A-1}{A}} \quad \text{and} \quad K \leq \frac{\log\log N}{2\log A} \tag{5.32}$$

then as N goes to infinity

$$\begin{aligned} &\| f_N^{(1,K)} - g_\alpha^{(1,K)} \|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \\ &\quad \leq \|\rho^0\|_{L^\infty}\, \varepsilon^{\frac{d-1}{d+1}} \exp\left(C\,(\log N)^{1/2}\, \log\log N\right). \end{aligned} \tag{5.33}$$

In particular, Estimate (5.21) follows from Proposition 5.8.

Proof We write

$$\begin{aligned} \| f_N^{(1,K)} - g_\alpha^{(1,K)} \|_{L^\infty} \leq \| f_N^{(1,K)} - \tilde f_{N,E,\delta}^{(1,K)} \|_{L^\infty} + \| g_\alpha^{(1,K)} - \tilde g_{\alpha,E,\delta}^{(1,K)} \|_{L^\infty} \\ + \| \tilde f_{N,E,\delta}^{(1,K)} - \tilde g_{\alpha,E,\delta}^{(1,K)} \|_{L^\infty}. \end{aligned}$$

Let $\bar a$, ε_0, δ, E satisfying (5.10). By gathering together the estimates in Propositions 5.4, 5.5, 5.6 and 5.7, we see that there exists C depending only on β and d such that, as N goes to infinity in the scaling $N\varepsilon^{d-1}\alpha^{-1} \equiv 1$, the following holds

$$\begin{aligned} &\| f_N^{(1,K)} - g_\alpha^{(1,K)} \|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \\ &\leq A^{K(K+1)} (C\alpha t)^{A^{K+1}} \left(e^{-\frac{\beta}{4}E^2} + \frac{A^{2(K+1)}\delta}{t} \right) \|\rho^0\|_{L^\infty} \\ &\quad + A^{(K+2)(K+1)} (C\alpha t)^{A^{K+1}} \left(E^d \left(\frac{\bar a}{\varepsilon_0}\right)^{d-1} + E^d \left(Et\right)^d \varepsilon_0^{d-1} \right. \\ &\qquad \left. + E \left(\frac{\varepsilon_0}{\delta}\right)^{d-1} \right) \|\rho^0\|_{L^\infty} \\ &\quad + A^{K(K+1)} (C\alpha t)^{A^{K+1}} \left(\frac{A^{2(K+1)}}{N} + \varepsilon\alpha \right) \|\rho^0\|_{L^\infty} \end{aligned}$$

with A, K introduced in Proposition 4.3.

To derive the upper bound (5.33), we choose for the parameters the following orders of magnitude

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

This leads to

$$\begin{aligned} &\left\| f_N^{(1,K)} - g_\alpha^{(1,K)} \right\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \\ &\quad \leq (C\,\alpha t)^{A^{K+1}} A^{2K(K+1)} \left(\varepsilon^{\frac{d-1}{d+1}} |\log \varepsilon|^d + \varepsilon^{d-1}\right) \|\rho^0\|_\infty \end{aligned}$$

from which (5.33) can be deduced in the scaling (5.32) since $A^K \leq \sqrt{\log N}$.

Equipped with all these estimates, we prove now Theorem 2.2.

Proof of Theorem 2.2 Propositions 4.3 and 5.8 imply that with the scaling (5.32)

$$\begin{aligned} &\left\| f_N^{(1)} - g_\alpha \right\|_{L^\infty([0,t]\times\mathbf{T}^d\times\mathbf{R}^d)} \\ &\quad \leq C\,\left(\gamma^A + C_0\,\varepsilon^{\frac{d-1}{d+1}} \exp\left(C\,(\log N)^{1/2}\,\log\log N\right)\right) \|\rho^0\|_{L^\infty} \\ &\quad \leq C\,\left(\frac{(\alpha t)^{A/(A-1)}}{\log\log N}\right)^A \|\rho^0\|_{L^\infty}, \end{aligned}$$

where we have used the relation $\gamma = \frac{(\alpha t)^{A/(A-1)}}{CK}$ of (4.13) with the choice $K = \lfloor \frac{\log\log N}{2\log A} \rfloor$. This conludes the proof of Theorem 2.2.

Note that the relevant scaling for this upper bound is $\alpha t = o\Big((\log\log N)^{(A-1)/A}\Big)$. □

6 Proof of the diffusive limit: proof of Theorem 2.3

In Theorem 2.2, we have shown that the tagged particle distribution $f_N^{(1)}(t,x,v)$ remains close to $M_\beta(v)\varphi_\alpha(t,x,v)$ where φ_α solves the linear-Boltzmann equation (1.3) on $\mathbf{T}^d \times \mathbf{R}^d$, with initial data $\rho^0(x)$. More generally, our proof implies that the whole trajectory of the tagged particle $\{x_1(s)\}_{s\leq T}$ can be approximated with high probability by the trajectory of $\{x_1^0(s)\}_{s\leq T}$ (see Lemma 5.3). 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 prove the diffusive limit. We first present an analytic approach to show that $\varphi_\alpha(\alpha\tau, x, v)$ can be approximated by the diffusion (2.11). Then we turn to an alternative method to show the convergence of the trajectory to a Brownian motion which will rely on probabilistic estimates for $\{x_1^0(\alpha\tau)\}_{\tau\le T}$.

In the following the macroscopic time variable will be denoted by $\tau \in [0, T]$.

6.1 Convergence to the heat equation

In this section we prove the result (2.12) stating the convergence of $f_N^{(1)}(\alpha\tau, x, v)$ to $M_\beta(v)\rho(\tau, x)$ where ρ solves the heat equation (2.11) on $\mathbf{T}^d$, with initial data $\rho^0(x)$. We show in Paragraph 6.1.1 that this can be reduced to proving that $\varphi_\alpha(\alpha\tau, x, v)$ can be approximated by a diffusion, which is a standard procedure (see [6]). For the sake of completeness, we recall the salient features of the proof in Paragraphs 6.1.2 and 6.1.3.

6.1.1 Approximation by the linear Boltzmann equation

The explicit convergence rate provided in Theorem 2.2 implies in particular that for any $\tau > 0$ and any $\alpha > 1$, in the limit $N \to \infty$, $N\varepsilon^{d-1}\alpha^{-1} = 1$, one has

$$\left\| f_N^{(1)}(\alpha\tau, x, v) - M_\beta(v)\varphi_\alpha(\alpha\tau, x, v)\right\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} \le C\left[\frac{\alpha^2\tau}{(\log\log N)^{\frac{A-1}{A}}}\right]^{\frac{A^2}{A-1}}, \tag{6.1}$$

where $A \ge 2$ can be taken arbitrarily large. It is therefore possible to take the limit $\alpha \to \infty$ while conserving a small right-hand side in (6.1), as soon as $\alpha \ll (\log\log N)^{\frac{A-1}{2A}}$.

Let us define

$$\widetilde{\varphi}_\alpha(\tau, x, v) := \varphi_\alpha(\alpha\tau, x, v),$$

which satisfies

$$\partial_\tau\widetilde{\varphi}_\alpha + \alpha\, v\cdot\nabla_x\widetilde{\varphi}_\alpha + \alpha^2\mathcal{L}\widetilde{\varphi}_\alpha = 0. \tag{6.2}$$

Then (2.12) follows directly from the following result

$$\sup_{\tau\in[0,T]}\ \sup_{(x,v)\in\mathbf{T}^d\times\mathbf{R}^d}\left|M_\beta(v)\big(\widetilde{\varphi}_\alpha(\tau, x, v) - \rho(\tau, x)\big)\right| \to 0 \tag{6.3}$$

in the limit $\alpha \to \infty$. The rest of this paragraph is devoted to the proof of (6.3). Notice that by the maximum principle on the heat equation, we may assume without loss of generality (up to regularizing ρ^0) that ρ^0 belongs to $C^4(\mathbf{T}^d)$, which will be useful at the end of the proof.

6.1.2 Hilbert's expansion

The formal Hilbert expansion consists in writing an asymptotic expansion of $\widetilde{\varphi}_\alpha$ in terms of powers of α^{-1}

$$\widetilde{\varphi}_\alpha(\tau, x, v) = \widetilde{\rho}_0(\tau, x, v) + \frac{1}{\alpha}\widetilde{\rho}_1(\tau, x, v) + \frac{1}{\alpha^2}\widetilde{\rho}_2(\tau, x, v) + \cdots,$$

in plugging that expansion in Eq. (6.2), and in canceling successively all the powers of α. This gives formally the following set of equations (where we have considered only the $O(1)$, $O(\alpha)$ and $O(\alpha^2)$ terms)

$$\begin{aligned} \mathcal{L}\widetilde{\rho}_0 &= 0, \\ v \cdot \nabla_x \widetilde{\rho}_0 + \mathcal{L}\widetilde{\rho}_1 &= 0, \\ \partial_\tau \widetilde{\rho}_0 + v \cdot \nabla_x \widetilde{\rho}_1 + \mathcal{L}\widetilde{\rho}_2 &= 0. \end{aligned} \tag{6.4}$$

In order to find the expressions for $\widetilde{\rho}_1$ and $\widetilde{\rho}_2$, as well as the equation on $\widetilde{\rho}_0$ (which we expect to be the heat equation), it is necessary to be able to invert the operator $\mathcal{L}$. This is made possible by the following result, whose proof can be found in [24] (in the case of the linearized Boltzmann equation, but it can easily be adapted to our situation). In the following, we define

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

The proof of the next result consists in noticing the decomposition $\mathcal{L} = a_\beta(v)\,\mathrm{Id} - \mathcal{K}$, where Id stands for the identity and $\mathcal{K}$ is a compact operator.

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

$$\left\{g \in L^2(\mathbf{R}^d, a_\beta M_\beta dv),\ \int_{\mathbf{R}^d} g(v)\, M_\beta(v)dv = 0\right\}.$$

Note that the first equation in (6.4) therefore reflects the fact that $\widetilde{\rho}_0$ does not depend on v.

We define the vector $b(v) = \left(b_k(v)\right)_{k\le d}$ with $\int_{\mathbf{R}^d} b(v)\, M_\beta(v)dv = 0$, by

$$\mathcal{L}b(v) = v. \tag{6.5}$$

Returning to (6.4), we have

$$\widetilde{\rho}_1(\tau, x, v) = \rho_1(\tau, x, v) + \overline{\rho}_1(\tau, x),$$

with

$$\rho_1(\tau, x, v) := -b(v) \cdot \nabla_x \widetilde{\rho}_0(\tau, x) \quad \text{and} \quad \overline{\rho}_1 \in \operatorname{Ker} \mathcal{L}.$$

Next we consider the last equation in (6.4) and we notice that for $\widetilde{\rho}_2$ to exist it is necessary for $\partial_\tau \widetilde{\rho}_0 + v \cdot \nabla_x \widetilde{\rho}_1$ to belong to the range of $\mathcal{L}$. Since $\widetilde{\rho}_0$ does not depend on v, this means that

$$\partial_\tau \widetilde{\rho}_0 + \int_{\mathbf{R}^d} v \cdot \nabla_x \widetilde{\rho}_1(\tau, x, v) M_\beta(v)\, dv = 0. \tag{6.6}$$

We then define the diffusion matrix $D(v) = \big(D_{k,\ell}(v)\big)_{k,\ell \le d}$, again with $\int_{\mathbf{R}^d} D_{k,\ell}(v)\, M_\beta(v) dv = 0$, by

$$\mathcal{L} D(v) := v \otimes b(v) - \int_{\mathbf{R}^d} v \otimes b(v)\, M_\beta(v) dv. \tag{6.7}$$

From the symmetry of the model, one can check (see [15] for instance) that there is a function γ such that

$$b(v) = \gamma(|v|) v.$$

Then an easy computation shows that $\widetilde{\rho}_0 = \rho$ where

$$\partial_\tau \rho - \kappa_\beta \Delta_x \rho = 0,$$

while the diffusion coefficient is given by

$$\kappa_\beta := \frac{1}{d} \int_{\mathbf{R}^d} v \mathcal{L}^{-1} v\ M_\beta(v) dv = \frac{1}{d} \int_{\mathbf{R}^d} \gamma(|v|) |v|^2\, M_\beta(v) dv, \tag{6.8}$$

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

$$\widetilde{\rho}_2(\tau, x, v) = \rho_2(\tau, x, v) + \overline{\rho}_2(\tau, x) - b(v) \cdot \nabla_x \overline{\rho}_1(\tau, x),$$

with

$$\rho_2(\tau, x, v) := D(v) : \operatorname{Hess} \rho(\tau, x) \quad \text{and} \quad \overline{\rho}_2 \in \operatorname{Ker} \mathcal{L}.$$

6.1.3 Proof of the convergence

Now let us prove (6.3). With the notation introduced in the previous paragraph, let us define

$$\Psi_\alpha(\tau, x, v) := \rho(\tau, x) + \frac{1}{\alpha}\rho_1(\tau, x, v) + \frac{1}{\alpha^2}\rho_2(\tau, x, v). \tag{6.9}$$

Then Ψ_α is almost a solution of (6.2): by construction one has

$$\partial_\tau \Psi_\alpha + \alpha\ v \cdot \nabla_x \Psi_\alpha + \alpha^2\ \mathcal{L}\Psi_\alpha = S_\alpha,$$

where the error term S_α is given by

$$S_\alpha(\tau, x, v) := \frac{1}{\alpha}\left(\partial_\tau \rho_1(\tau, x, v) + v \cdot \nabla_x \rho_2(\tau, y, v) + \frac{1}{\alpha}\partial_\tau \rho_2(\tau, y, v)\right). \tag{6.10}$$

Defining

$$R_\alpha(\tau, x, v) := \Psi_\alpha(\tau, x, v) - \widetilde{\varphi}_\alpha(\tau, x, v)$$

we have thanks to (6.2)

$$\partial_\tau R_\alpha + \alpha\ v \cdot \nabla_x R_\alpha + \alpha^2\ \mathcal{L}R_\alpha = S_\alpha$$

and the result (6.3) then follows from the maximum principle which states that

$$\begin{aligned}\|M_\beta R_\alpha\|_{L^\infty([0,T]\times\mathbf{T}^d\times\mathbf{R}^d)} \leq C(T)\big(&\|M_\beta R_\alpha(0)\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)} \\ &+\|M_\beta S_\alpha\|_{L^\infty([0,T]\times\mathbf{T}^d\times\mathbf{R}^d)}\big).\end{aligned}$$

We note that S_α involves spatial derivatives of ρ of order at most 4, thus from the maximum principle for the heat equation, each term of $M_\beta S_\alpha$ is bounded in L^∞ norm by α^{-1}. The same clearly holds for the initial data $M_\beta R_\alpha(0, x, v)$ since

$$R_\alpha(0, x, v) = \Psi_\alpha(0, x, v) - \widetilde{\varphi}_\alpha(0, x, v) = \frac{1}{\alpha}\rho_1(0, x, v) + \frac{1}{\alpha^2}\rho_2(0, x, v).$$

It follows that

$$\|M_\beta(\Psi_\alpha - \widetilde{\varphi}_\alpha)\|_{L^\infty([0,T]\times\mathbf{T}^d\times\mathbf{R}^d)} \leq \frac{C(T)}{\alpha}$$

and thanks to (6.9), the convergence result (6.3) is proved.

Remark 6.2 We have considered here the case when $\rho^0 = \rho^0(x)$. In the case of ill-prepared initial data, namely if $\rho^0 = \varphi^0(x, v)$, then the same analysis works provided the following ansatz is used

$$\Psi_\alpha(\tau, x, v) := \rho(\tau, x) + \frac{1}{\alpha}\rho_1(\tau, x, v) + \frac{1}{\alpha^2}\rho_2(\tau, x, v) + \Pi^\perp\left(e^{-\alpha\tau\mathcal{L}}\varphi^0\right),$$

where $\Pi^\perp$ is the orthogonal projector onto $(\operatorname{Ker}\mathcal{L})^\perp$.

6.2 Convergence to the Brownian motion

Let us denote the tagged particle by

$$\Xi(\tau) := x_1(\alpha\tau).$$

In the following, $\mathbb{E}_N, \mathbb{P}_N$ will refer to its expectation and probability with respect to the initial data sampled from the density f_N^0. To prove the convergence of the tagged particle to a Brownian motion, one needs to check (see [8], Chapter 2)

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

$$\lim_{N\to\infty} \mathbb{E}_N\Big(h_1\big(\Xi(\tau_1)\big)\dots h_\ell\big(\Xi(\tau_\ell)\big)\Big) = \mathbb{E}\Big(h_1\big(B(\tau_1)\big)\dots h_\ell\big(B(\tau_\ell)\big)\Big), \tag{6.11}$$

  where $\{h_1, \dots, h_\ell\}$ is a collection of continuous functions in $\mathbf{T}^d$. Notice that these marginals refer to time averages and not to the number of particles.
- The tightness of the sequence, that is for any $\tau \in [0, T]$

$$\forall \xi > 0, \qquad \lim_{\eta\to 0}\lim_{N\to\infty} \mathbb{P}_N\left(\sup_{\tau<\sigma<\tau+\eta}\big|\Xi(\sigma) - \Xi(\tau)\big| \geq \xi\right) = 0. \tag{6.12}$$

Note that (6.11) requires to understand time correlations and thus we are going to adapt Theorem 2.2 to this new framework.

Step 1. *Finite dimensional marginals*. First, we are going to rewrite the time correlations in terms of collision trees. A similar approach was devised in Lebowitz, Spohn [30] to derive an information on the true particle trajectories (in the physical space) from the Duhamel series. Let $t_1 < \cdots < t_\ell$ be an increasing collection of times and $H_\ell = \{h_1, \dots, h_\ell\}$ a collection of ℓ smooth functions. Define the biased distribution at time $t > t_\ell$ as follows

$$\begin{aligned}
&\int_{\mathbf{T}^{Nd}\times\mathbf{R}^{Nd}} dZ_N f_{N,H_\ell}(t,Z_N)\Phi(Z_N) \\
&\quad := \mathbb{E}_N\Big(h_1\big(x_1(t_1)\big)\dots h_\ell\big(x_1(t_\ell)\big)\,\Phi\big(Z_N(t)\big)\Big) \\
&\quad = \int_{\mathbf{T}^{Nd}\times\mathbf{R}^{Nd}} dZ_N\, f_N^0(Z_N)\, h_1\big(x_1(t_1)\big)\dots h_\ell\big(x_1(t_\ell)\big)\,\Phi\big(Z_N(t)\big), \quad (6.13)
\end{aligned}$$

for any test function Φ. We stress that by construction the biased distribution $f_{N,H_\ell}(t,Z_N)$

- Is in general no longer normalized by 1
- Is symmetric with respect to the $N-1$ last variables.

The corresponding marginals are

$$f^{(s)}_{N,H_\ell}(t,Z_s) := \int f_{N,H_\ell}(t,Z_N)\,dz_{s+1}\dots dz_N. \quad (6.14)$$

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 (3.1) for $s < N$. Applying the iterated Duhamel formula (3.5), we get

$$f^{(s)}_{N,H_\ell}(t) = \sum_{m=0}^{N-s} Q_{s,s+m}(t-t_\ell)\,f^{(s+m)}_{N,H_\ell}(t_\ell). \quad (6.15)$$

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 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 with the motion of the tagged particle

$$\begin{aligned}
f^{(1)}_{N,H_\ell}(t) = \sum_{m_1+\dots+m_\ell=0}^{N-1} & Q_{1,1+m_1}(t-t_\ell)\Big(h_\ell Q_{1+m_1,1+m_2}(t_\ell - t_{\ell-1})\Big(h_{\ell-1}\dots \\
& Q_{1+m_1+\dots+m_{\ell-1},1+m_1+\dots+m_\ell}(t_1)\Big)\,f_N^{(1+m_1+\dots+m_\ell)}(0). \quad (6.16)
\end{aligned}$$

This identity holds for any N and any time.

In order to check (6.12), we need also to generalize the identity to consider correlations of the form

$$\mathbb{E}_N\Big(h\big(x_1(t_1)-x_1(s)\big)\dots h\big(x_1(t_\ell)-x_1(s)\big)\Big) \quad (6.17)$$

for a smooth function h with $s < t_1 < \cdots < t_\ell$. 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 (6.16) for any accuracy $\xi > 0$ of the approximation. Thus (6.17) can be computed in terms of collision trees which are now weighted by the factor $h\big(x_1(t_1) - x_1(s)\big) \dots h\big(x_1(t_\ell) - x_1(s)\big)$.

Step 2. *The limit process*. In the Boltzmann Grad limit, the memory of the system is lost and the tagged particle behavior becomes equivalent to a Markov process. We define

$$\bar{x}_1(t) = \bar{x}_1(0) + \int_0^t \bar{v}_1(s)\, ds \tag{6.18}$$

as an additive functional of the Markov chain $\{\bar{v}_1(s)\}_{s \geq 0}$ with generator $\alpha \mathcal{L}$ introduced in (1.3). Initially $(\bar{x}_1(0), \bar{v}_1(0))$ is distributed according to $\rho^0(x) M_\beta(v)$. The expectation associated to this Markov chain is denoted by $\mathbb{E}_{M_\beta}$.

Let $t_1 < \cdots < t_\ell$ be an increasing collection of times and $H_\ell = \{h_1, \dots, h_\ell\}$ a collection of ℓ smooth functions. As in (6.13), we define $g_{\alpha, H_\ell}(t)$ as the biased distribution of the Markov chain $\bar{z}_1(t) = \big(\bar{x}_1(t), \bar{v}_1(t)\big)$

$$\forall t \in [t_k, t_{k+1}],$$
$$\int_{\mathbf{T}^d \times \mathbf{R}^d} g_{\alpha, H_\ell}(t, z)\, \Phi(z)\, dz = \mathbb{E}_{M_\beta}\Big(h_1\big(\bar{x}_1(t_1)\big) \dots h_\ell\big(\bar{x}_1(t_\ell)\big)\, \Phi\big(\bar{z}_1(t)\big)\Big),$$

with $t_{\ell+1} = \infty$. One can consider a measure [cf. (3.11)] including as well the background density of an ideal gas. The marginals of this measure are

$$g_{\alpha, H_\ell}^{(s)}(t, Z_s) = g_{\alpha, H_\ell}(t, z_1) \prod_{i=2}^{s} M_\beta(v_i). \tag{6.19}$$

As in (6.16), the distribution can be rewritten in terms of a Duhamel series

$$g_{\alpha, H_\ell}(t) = \sum_{m_1 + \cdots + m_\ell = 0}^{\infty} Q^0_{1, 1+m_1}(t - t_\ell)\Big(h_\ell Q^0_{1+m_1, 1+m_2}(t_\ell - t_{\ell-1})\Big(h_{\ell-1} \dots$$
$$Q^0_{1+m_1+\cdots+m_{\ell-1}, 1+m_1+\cdots+m_\ell}(t_1)\Big) g_{\alpha, H_\ell}^{(1+m_1+\cdots+m_\ell)}(0). \tag{6.20}$$

This representation allows us to rephrase the Markov chain expectations in terms of the Boltzmann hierarchy. In this series, a lot of cancelations occur

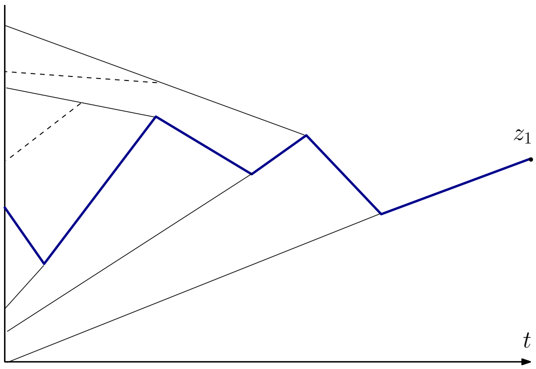

Fig. 5 A collision tree for the Boltzmann hierarchy is depicted. The path of z_1 is the backbone of the tree with branchings at each new collision. There cannot be further branches as any collision with a new particle would lead to a cancellation in the collision operator. Thus the trees involving the branches in *dashed line* do not contribute to the Duhamel series in the Boltzmann hierarchy

(see Fig. 5). Indeed, the only relevant collision trees are made of a single backbone (the trajectory of z_1^0) with branches representing the collisions of z_1^0 with the ideal gas, but no further ramification. In step 3, we shall not use these cancellations and simply compare the series (6.16) and (6.20) term by term in order to show that in the Boltzmann Grad limit when $\alpha \ll \sqrt{\log\log N}$

$$\lim_{\alpha\to\infty} \mathbb{E}_{M_\beta}\Big(h_1\big(\bar{x}_1(\alpha\tau_1)\big)\dots h_\ell\big(\bar{x}_1(\alpha\tau_\ell)\big)\Big) - \mathbb{E}_N\Big(h_1\big(x_1(\alpha\tau_1)\big)\dots h_\ell\big(x_1(\alpha\tau_\ell)\big)\Big) = 0. \tag{6.21}$$

As $\mathcal{L}$ has a spectral gap, the invariance principle holds for the position of the Markov process $\bar{x}_1$ (see [27] 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_β [see (6.8)], i.e. that for any smooth functions $\{h_1, \dots, h_\ell\}$ defined in $\mathbf{T}^d$,

$$\lim_{\alpha\to\infty} \mathbb{E}_{M_\beta}\Big(h_1\big(\bar{x}_1(\alpha\tau_1)\big)\dots h_\ell\big(\bar{x}_1(\alpha\tau_\ell)\big)\Big) = \mathbb{E}\Big(h_1\big(B(\tau_1)\big)\dots h_\ell\big(B(\tau_\ell)\big)\Big). \tag{6.22}$$

The diffusion coefficient κ_β defined in (6.8) can be interpreted in terms of the variance of the position properly rescaled in time (see [27] page 47).

Step 3. *Approximation of the finite dimensional marginals*. We turn now to the proof of (6.21) which combined with (6.22) will show the convergence (6.11) of the marginals of the tagged particle sampled at different times.

Suppose now that the collection H_ℓ satisfies the uniform bounds on $\mathbf{T}^d$

$$\forall i \le \ell, \qquad 0 \le h_i(x_1) \le m. \tag{6.23}$$

Thus the $f^{(s)}_{N,H_\ell}$ satisfy the maximum principle (4.2) with an extra factor m^ℓ. The pruning procedure on the collision trees therefore applies also in this case and enables us to restrict to trees with at most A^{K+1} collisions during the time interval $[0, t]$. Furthermore, the comparison of the trajectories for $f^{(1)}_{N,H_\ell}$ and g_{α,H_ℓ} can be achieved in the same way as before on a tree with less than A^{K+1} 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 $h_1\big(x^0_1(\alpha\tau_1)\big)\dots h_\ell\big(x^0_1(\alpha\tau_\ell)\big)$. We recall that the pseudo-trajectories of x_1 and x^0_1 coincide at any time, thus bounds similar to (2.9) hold

$$\begin{aligned}&\big\| f^{(1)}_{N,H_\ell}(\alpha\tau, y, v) - g_{\alpha,H_\ell}(\alpha\tau, y, v)\big\|_{L^\infty(\mathbf{T}^d\times\mathbf{R}^d)}\\&\quad\le Cm^\ell\,\|\rho^0\|_\infty\left[\frac{\tau\alpha^2}{(\log\log N)^{\frac{A-1}{A}}}\right]^{\frac{A^2}{A-1}}.\end{aligned}\tag{6.24}$$

This implies (6.21), hence (6.11) thanks to (6.22).

Step 4. *Tightness*. In order to evaluate (6.12), it is enough to sample the trajectory of the tagged particle at the times $\tau_i = \{\tau + \frac{u_N}{\alpha}\, i\}_{i\le\ell_N}$ for $u_N = \frac{1}{\log N}$ and with $\ell_N := \alpha\eta/u_N$. Indeed, we can decompose the path deviations into two terms

$$\begin{aligned}&\mathbb{P}_N\left(\sup_{\tau<\sigma<\tau+\eta}\big|\Xi(\sigma)-\Xi(\tau)\big|\ge 2\xi\right)\\&\quad\le 1-\mathbb{P}_N\left(\bigcap_{i=1}^{\ell_N}\big\{\big|\Xi(\tau_i)-\Xi(\tau)\big|<\xi\big\}\right)\\&\qquad+\sum_{i=1}^{\ell_N}\mathbb{P}_N\left(\sup_{\tau_i<\sigma<\tau_{i+1}}\big|\Xi(\sigma)-\Xi(\tau_i)\big|\ge\xi\right).\end{aligned}\tag{6.25}$$

We shall first evaluate the last term in the right-hand side which involves only events occurring in a microscopic time scale of length u_N. Given $i\le\ell_N$, let $t_i := iu_N + \alpha\tau$ and $t_{i+1} := u_N + t_i$ then

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

In order to control the tagged particle fluctuations, it is enough to bound its velocity in the time interval $[t_i, t_{i+1}]$

$$\mathbb{P}_N\left(\sup_{t_i<s<t_{i+1}}\Big|\int_{t_i}^{s} v_1(s')ds'\Big|\geq\xi\right)\leq\mathbb{P}_N\left(\int_{t_i}^{t_{i+1}}\big|v_1(s')\big|ds'\geq\xi\right)$$
$$\leq\|\rho^0\|_\infty\,\widehat{\mathbb{P}}_N\left(\int_0^{u_N}\big|v_1(s')\big|ds'\geq\xi\right),$$

where we used the maximum principle in the last inequality and $\widehat{\mathbb{P}}_N$ denotes the dynamics starting from the invariant measure $M_{N,\beta}$. Following the strategy in [2] to bound this probability, we write

$$\widehat{\mathbb{P}}_N\left(\int_0^{u_N}\big|v_1(s')\big|ds'\geq\xi\right)\leq\exp\left(-\frac{\xi}{u_N}\right)\widehat{\mathbb{E}}_N\left(\exp\left(\frac{1}{u_N}\int_0^{u_N}\big|v_1(s')\big|ds'\right)\right).$$

Using Jensen's inequality and the invariant measure, we get

$$\widehat{\mathbb{E}}_N\left(\exp\left(\frac{1}{u_N}\int_0^{u_N}\big|v_1(s')\big|ds'\right)\right)\leq\frac{1}{u_N}\int_0^{u_N}ds'\,\widehat{\mathbb{E}}_N\left(\exp\left(\big|v_1(s')\big|\right)\right)$$
$$=\mathbb{E}_{M_{N,\beta}}\left(\exp\left(\big|v_1\big|\right)\right)\leq c_\beta,$$

where c_β is a constant depending only on β. Since $u_N=\frac{1}{\log N}$, we have shown that for any $\xi>0$, the probability of a deviation in a very short time vanishes when N goes to infinity

$$\sum_{i=1}^{\ell_N}\mathbb{P}_N\left(\sup_{\tau_i<\sigma<\tau_{i+1}}\big|\Xi(\sigma)-\Xi(\tau_i)\big|\geq\xi\right)\leq c_\beta\,\ell_N\,\exp\left(-\xi\,\log N\right)\xrightarrow[N\to\infty]{}0.$$

The tightness for the process $\bar{x}_1$ derived in [27] (Theorem 2.32 page 74) implies that for any $\xi>0$ and $\ell_N=\alpha\eta/u_N$

$$\lim_{\eta\to0}\lim_{N\to\infty}\mathbb{P}_{M_\beta}\left(\bigcap_{i=1}^{\ell_N}\left\{\big|\bar{x}_1(\alpha\tau_i)-\bar{x}_1(\alpha\tau)\big|<\xi/2\right\}\right)=1.$$

By comparison, we are going to show that the same result holds also for the tagged particle x_1. Using (6.25), this will complete the proof of (6.12).

Let $h_\xi(w)=1_{\{|w|\leq\xi/2\}}$, then it is enough to show that

$$\left|\mathbb{E}_N\left(\prod_{i=1}^{\ell_N}h_\xi\left(x_1(\alpha\tau_i)-x_1(\alpha\tau)\right)\right)-\mathbb{E}_{M_\beta}\left(\prod_{i=1}^{\ell_N}h_\xi\left(\bar{x}_1(\alpha\tau_i)-\bar{x}_1(\alpha\tau)\right)\right)\right|$$
$$\leq C\|\rho^0\|_\infty\left[\frac{\tau\alpha^2}{(\log\log N)^{\frac{A-1}{A}}}\right]^{\frac{A^2}{A-1}}. \tag{6.26}$$

At this stage, it is enough to use the fact that probabilities of the form (6.17) can also be evaluated in terms of weighted trees as in Step 1. Since h_ξ 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. For good pseudo-trajectories, the paths of x_1 and x_1^0 coincide, therefore modifying the Duhamel series by $\prod_{i=1}^{\ell_N} h_\xi\big(x_1(\alpha\tau_i) - x_1(\alpha\tau)\big)$ does not alter the comparison established in Proposition 5.7. This concludes the proof of tightness.

Acknowledgments The authors are grateful to Laurent Desvillettes, François Golse, Herbert Spohn and Balint Toth for fruitful discussions and for suggesting interesting bibliographical references. Finally we extend our thanks to the anonymous referees for suggesting many improvements in the manuscript. The work of I. Gallagher has been supported by the grant ANR-12-BS01-0013-01.

7 Appendix A: Asymptotic control of the exclusion

For the sake of completeness, we recall here the proof of Proposition 3.2. We omit all subscripts β to simplify the presentation.

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

We first prove that in the scaling $N\varepsilon^{d-1} \equiv \alpha$, with $\alpha \ll 1/\varepsilon$,

$$1 \le \mathcal{Z}_N^{-1}\mathcal{Z}_{N-s} \le \big(1 - \varepsilon\alpha\kappa_d\big)^{-s}, \tag{7.1}$$

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

$$\mathcal{Z}_N \le \mathcal{Z}_{N-s}.$$

Let us prove the second inequality. We have by definition

$$\mathcal{Z}_{s+1} = \int_{\mathbf{T}^{d(s+1)}} \left(\prod_{1\le i\ne j\le s+1} \mathbf{1}_{|x_i-x_j|>\varepsilon} \right) dX_{s+1}.$$

By Fubini's equality, we deduce

$$\mathcal{Z}_{s+1} = \int_{\mathbf{T}^{ds}} \left(\int_{\mathbf{T}^d} \left(\prod_{1\le i\le s} \mathbf{1}_{|x_i-x_{s+1}|>\varepsilon} \right) dx_{s+1} \right) \left(\prod_{1\le i\ne j\le s} \mathbf{1}_{|x_i-x_j|>\varepsilon} \right) dX_s.$$

Since

$$\int_{\mathbf{T}^d} \left(\prod_{1 \le i \le s} \mathbf{1}_{|x_i - x_{s+1}| > \varepsilon} \right) dx_{s+1} \ge 1 - \kappa_d s \varepsilon^d ,$$

we deduce the lower bound

$$\mathcal{Z}_{s+1} \ge \mathcal{Z}_s (1 - \kappa_d s \varepsilon^d) \ge \mathcal{Z}_s (1 - \kappa_d \varepsilon \alpha),$$

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

$$\mathcal{Z}_N \ge \mathcal{Z}_{N-s} \big(1 - \varepsilon \alpha \kappa_d\big)^s .$$

That proves (7.1).

- *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 $s \le N$

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

We deduce, by symmetry,

$$M_N^{(s)} = \mathcal{Z}_N^{-1} \mathbf{1}_{Z_s \in \mathcal{D}_\varepsilon^s} M^{\otimes s} \left(\mathcal{Z}_{N-s} - \mathcal{Z}^\flat_{(s+1,N)} \right) \tag{7.2}$$

with the notation

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

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

$$\mathbf{1}_{Z_s \in \mathcal{D}^s_\varepsilon} M^{\otimes s} - M_N^{(s)} = \Big(1 - \mathcal{Z}_N^{-1} \mathcal{Z}_{N-s}\Big) \mathbf{1}_{Z_s \in \mathcal{D}^s_\varepsilon} M^{\otimes s} + \mathcal{Z}_N^{-1} \mathcal{Z}^\flat_{(s+1,N)} \mathbf{1}_{Z_s \in \mathcal{D}^s_\varepsilon} M^{\otimes s}. \tag{7.3}$$

By (7.1), there holds $1 - \mathcal{Z}_N^{-1} \mathcal{Z}_{N-s} \to 0$ 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 (7.3) tends to 0 as N goes to ∞. Besides, by

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

we bound

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

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

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

so that

$$\mathcal{Z}^\flat_{(s+1,N)} \le s(N-s)\varepsilon^d \kappa_d \mathcal{Z}_{N-s-1}. \tag{7.4}$$

By (7.1), we obtain

$$\mathcal{Z}_N^{-1} \mathcal{Z}^\flat_{(s+1,N)} \le \varepsilon \alpha s \kappa_d \big(1 - \varepsilon \alpha \kappa_d\big)^{-(s+1)},$$

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 (7.3).

This completes the proof of Proposition 3.2.

8 Appendix B: Recollisions in the torus

We show here how to adapt the arguments of [21] to prove Lemma 5.2.

- 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 $u \in [0, t]$ such that

$$d\big((x_1 - v_1 u), (x_2 - v_2 u)\big) \leq \varepsilon,$$

boils down to having

$$(x_1 - v_1 u) - (x_2 - v_2 u) \in \bigcup_{k \in \mathbf{Z}^d} B_\varepsilon(k).$$

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

$$(x_1^0 - v_1 u) - (x_2^0 - v_2 u) \in \bigcup_{k \in \mathbf{Z}^d} B_{3\bar{a}}(k).$$

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

$$s(v_1 - v_2) \in \left(\bigcup_{k \in \mathbf{Z}^d} B_{3\bar{a}}(x_1^0 - x_2^0 + k) \right) \cap B_0(2Et).$$

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 $(4Et)^d$) are of solid angle $c\,\bar{a}^{d-1}$.

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

$$|K(\bar{x}_1 - \bar{x}_2, \varepsilon_0, \bar{a})| \leq C E^d \left(\left(\frac{\bar{a}}{\varepsilon_0} \right)^{d-1} + (Et)^d \bar{a}^{d-1} \right).$$

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

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

boils down to having

$$u(v_1 - v_2) \in B_{3\varepsilon_0}(x_1^0 - x_2^0 + k), \tag{8.1}$$

for some $k \in \mathbf{Z}^d \cap B_{2Et}(x_2^0 - x_1^0)$.

- If $|x_1^0 - x_2^0 + k| \geq 1/4$, condition (8.1) implies that $v_1 - v_2$ belongs to the intersection of $B_{2E}(0)$ and some cone of vertex 0 and solid angle ε_0^{d-1}.
- If $|x_1^0 - x_2^0 + k| \leq 1/4$ (which can happen only for one value of k), denoting by n any unit vector normal to $\bar{x}_1 - \bar{x}_2 + k$, we deduce from (8.1) that

$$u|(v_1 - v_2) \cdot n| \leq 3\varepsilon_0$$

  from which we deduce that $v_1 - v_2$ belongs to the intersection of $B_{2E}(0)$ and some cylinder of radius ε_0/δ.

The union $K_\delta(x_1^0 - x_2^0, \varepsilon_0, \bar{a})$ of these "bad" sets is therefore of size

$$|K_\delta(x_1^0 - x_2^0, \varepsilon_0, \bar{a})| \leq CE\left(\left(\frac{\varepsilon_0}{\delta}\right)^{d-1} + E^{d-1}\big(Et\big)^d {\varepsilon_0}^{d-1}\right).$$

The lemma is proved. □

References

1. Alexander, R.: The infinite hard sphere system, Ph.D. dissertation, Department of Mathematics, University of California, Berkeley (1975)
2. Alexander, R.: Time evolution for infinitely many hard spheres. Comm. Math. Phys. **49**(3), 217–232 (1976)
3. Bardos, C.: Problèmes aux limites pour les quations aux dérivées partielles du premier ordre à coefficients réels; théorèmes d'approximation; application à l'équation de transport. Ann. Sci. École Norm. Sup. **4**(3), 185–233 (1970)
4. Bardos, C., Golse, F., Levermore, C.D.: Fluid dynamic limits of the Boltzmann equation I. J. Stat. Phys. **63**, 323–344 (1991)
5. Bardos, C., Golse, F., Levermore, C.D.: Fluid dynamic limits of kinetic equations II: convergence proofs for the Boltzmann equation. Comm. Pure Appl. Math. **46**, 667–753 (1993)
6. Bardos, C., Santos, R., Sentis, R.: Diffusion approximation and computation of the critical size. Trans. Am. Math. Soc **284**(2), 617–649 (1984)
7. van Beijeren, H., Lanford III, O.E., Lebowitz, J.L., Spohn, H.: Equilibrium time correlation functions in the low density limit. J. Stat. Phys. **22**, 237–257 (1980)
8. Billingsley, P.: Probability and measure. Wiley series in probability and mathematical statistics. Wiley (1995)
9. Boldrighini, C., Bunimovich, L.A., Sinai, Y.: On the Boltzmann equation for the Lorentz gas. J. Stat. Phys. **32**(3), 477–501 (1983)
10. Bunimovich, L.A., Sinai, Y.G.: Statistical properties of Lorentz gas with periodic configuration of scatterers. Comm. Math. Phys. **78**, 479–497 (1980/81)
11. Caglioti, E., Golse, F.: On the distribution of free path lengths for the periodic Lorentz gas III. Commun. Math. Phys. **236**, 199–221 (2003)
12. Caglioti, E., Golse, F.: On the Boltzmann-Grad limit for the two dimensional periodic Lorentz gas. J. Stat. Phys. **141**, 264–317 (2010)
13. Cercignani, C., Illner, R., Pulvirenti, M.: The mathematical theory of dilute gases. Springer, New York (1994)

14. De Roeck, W., Fröhlich, J.: Diffusion of a massive quantum particle coupled to a quasi-free thermal medium. Comm. Math. Phys. **303**(3), 613–707 (2011)
15. Desvillettes, L., Golse, F.: A remark concerning the Chapman-Enskog asymptotics. In: Perthame, B. (ed.) Advances in kinetic theory and computing. Ser. Adv. Math. Appl. Sci., vol. 22, pp. 191–203. World Science Publishing, River Edge (1994)
16. Desvillettes, L., Pulvirenti, M.: The linear Boltzmann equation for long-range forces: a derivation from particles. Math. Meth. Mod. Appl. Sci. **9**(8), 11231145 (1999)
17. Desvillettes, L., Ricci, V.: A rigorous derivation of a linear kinetic equation of Fokker–Planck type in the limit of grazing collisions. J. Stat. Phys. **104**(5–6), 1173–1189 (2001)
18. Erdős, L.: Lecture notes on quantum Brownian motion, quantum theory from small to large scales: lecture notes of the Les Houches summer school, 95, Oxford University Press (2010)
19. Erdős, L., Salmhofer, M., Yau, H.T.: Quantum diffusion of the random Schrodinger evolution in the scaling limit. Acta Math. **200**(2), 211–278 (2008)
20. Esposito, R., Marra, R., Yau, H.T.: Navier–Stokes equations for stochastic particle systems on the lattice. Comm. Math. Phys. **182**, 395–456 (1996)
21. Gallagher, I., Saint-Raymond, L., Texier, B.: From Newton to Boltzmann: the case of hard-spheres and short-range potentials, Zürich lectures in advanced mathematics 18. Erratum to Chapter 5 (2014)
22. Gallavotti, G.: Statistical mechanics. A short treatise. Texts and monographs in physics. Springer, Berlin (1999)
23. Golse, F.: On the periodic Lorentz gas and the Lorentz kinetic equation. Ann. Fac. Sci. Toulouse Math. **17**, 735–749 (2008)
24. Hilbert, D.: Begründung der kinetischen Gastheorie. (German). Math. Ann. **72**(4), 562–577 (1912)
25. Kac, M.: Probability and related topics in physical sciences, Am. Math. Soc., p. 1 (1959)
26. King, F.: BBGKY hierarchy for positive potentials, Ph.D. dissertation, Dept. Mathematics, Univ. California, Berkeley (1975)
27. Komorowski, T., Landim, C., Olla, S.: Fluctuations in Markov processes. Time symmetry and martingale approximation. Grundlehren der Mathematischen Wissenschaften, vol. 345. Springer, Heidelberg (2012)
28. Lanford, O.E.: Time evolution of large classical systems. In: Moser, J. (ed.) Lecture notes in physics, vol. 38, pp. 1–111, Springer (1975)
29. Lebowitz, J., Spohn, H.: Microscopic basis for Fick's law for self-diffusion. J. Stat. Phys. **28**, 539–556 (1982)
30. Lebowitz, J., Spohn, H.: Steady state self-diffusion at low density. J. Stat. Phys. **29**, 39–55 (1982)
31. Lorentz, H.: Le mouvement des électrons dans les métaux. Arch. Neerl. **10**, 336–371 (1905)
32. Marklof, J.: Kinetic transport in crystals. In: Proceedings of the XVI International Congress on Mathematical Physics, Prague 2009, World Scientific, pp. 162–179 (2010)
33. Marklof, J., Strömbergsson, A.: The Boltzmann–Grad limit of the periodic Lorentz gas. Ann. Math. **174**, 225–298 (2011)
34. Marklof, J., Toth, B.: Superdiffusion in the periodic Lorentz gas, arXiv:1403.6024, preprint (2014)
35. Pettersson, R.: On weak and strong convergence to equilibrium for solutions to the linear Boltzmann equation. J. Stat. Phys. **72**, 355–380 (1993)
36. Olla, S., Varadhan, S., Yau, H.-T.: Hydrodynamical limit for a Hamiltonian system with weak noise. Commun. Math. Phys. **155**, 523–560 (1993)
37. Pulvirenti, M., Saffirio, C., Simonella, S.: On the validity of the Boltzmann equation for short range potentials. Rev. Math. Phys. **26**(2), 1450001 (2014)
38. Quastel, J., Yau, H.-T.: Lattice gases, large deviations, and the incompressible Navier–Stokes equations. Ann. Math. **148**, 51–108 (1998)

39. Saint-Raymond, L.: Hydrodynamic limits of the Boltzmann equation. Lecture notes in mathematics, vol. 1971, Springer (2009)
40. Simonella, S.: Evolution of correlation functions in the hard sphere dynamics. J. Stat. Phys. **155**(6), 1191–1221 (2014)
41. Spohn, H.: The Lorentz process converges to a random flight process. Commun. Math. Phys. **60**, 277–290 (1978)
42. Spohn, H.: Large scale dynamics of interacting particles, vol. 174, Springer (1991)
43. Szasz, D., Toth, B.: Towards a unified dynamical theory of the Brownian particle in an ideal gas. Comm. Math. Phys. **111**, 41–62 (1987)
44. Uchiyama, K.: Derivation of the Boltzmann equation from particle dynamics. Hiroshima Math. J. **18**(2), 245–297 (1988)