LONG TIME VALIDITY OF THE LINEARIZED BOLTZMANN UNCUT-OFF AND THE LINEARIZED LANDAU EQUATIONS FROM THE NEWTON LAW

CORENTIN LE BIHAN

Abstract. We provide a rigorous justification of the linearized Boltzmann and Landau equations for interacting particle systems with long-range interaction. The result shows that for a system of N Hamiltonian particles governed by truncated power law potentials of the form $\mathscr{U}_{\varepsilon_{\text{eff}}}(r) \sim |r/\varepsilon_{\text{eff}}|^{-s}$ near $r \approx 0$ (with ε_{eff} the effective radius of the particles), the covariance of the equilibrium fluctuations converges to solutions of kinetic equations in appropriate scaling limits $\varepsilon_{\text{eff}} \to 0$ and $N \to \infty$, corresponding to a low density regime . We prove that in dimension 3 , for $s \in (1, \infty)$, the limiting system approaches the uncutoff linearized Boltzmann equation for the scaling $N{\varepsilon_{\text{eff}}}^2 = 1$. The Coulomb singularity $s = 1$ appears as a threshold value. Kinetic scaling limits with $s \in (0, 1]$ universally converge to the linearized Landau equation, and we prove the onset of the Coulomb logarithm for $s = 1$.

Contents

1. Introduction 2
1.1. Modification of the scaling parameters and the sub-Coulomb case ($s \in [0, 1)$) 4
1.2. State of the art 5
2. Definition of the system and strategy of the proof 6
2.1. The Hamiltonian dynamics 6
2.2. Grand-canonical ensemble and stationary measure 6
2.3. Binary collision, scattering, and definition of the linearized Boltzmann operator 7
2.4. Convergence to the linearized Boltzmann equation with a cut-off 9
2.5. Derivation of the linearized Landau equation and Boltzmann equation without cut-off 9
2.6. Strategy of the proof of Theorem 1 11
3. Development along pseudotrajectories and time sampling 15
3.1. Dynamical cluster development 15
3.2. Conditioning 19
3.3. The main part of the cumulant 19
3.4. Iteration of the pseudotrajectory development 21
3.5. The decomposition of the covariance 23
4. Quasi-orthogonality estimates 26
5. Clustering estimations without recollision 31
6. Treatment of the main part 35
6.1. Duality formula 35
6.2. Linearized Boltzmann equation 39
7. Estimation of non-pathological recollisions 42
8. Estimation of the local recollisions 44
Appendix A. The linearized Boltzmann operator without cut-off 53
Appendix B. Geometrical estimates 54
B.1. Estimation the length scattering time 54
B.2. Proof of Proposition 7.2 54
B.3. Proof of (3.6) 59
B.4. Proof of (8.10) 59

E-mail address: corentin.le.bihan@ulb.be.

Key words and phrases. Low density, Boltzmann-Grad limit, Grazing collision limit, Boltzmann equation, Landau equation, Kinetic theory.

1. Introduction

In kinetic theory, a gas of particles can be modeled by a large system of N classical particles interacting *via* a potential $\mathscr{U}(\cdot/\varepsilon_{\text{eff}})$, with ε_{eff} the *effective radius* of a particle . The dynamics is Hamiltonian and obeys the Newton laws with assigned potential $\mathscr{U}$. In dimension 3, the power laws $\mathscr{U}_s(r) := r^{-s}$, $s \geq 1$ play a fundamental role, in particular the Coulomb case $s = 1$. One of the goals of kinetic theory is the description of such a gas in the limit $N \to \infty$, $\varepsilon_{\text{eff}} \to 0$. Of course, the asymptotic behavior depends on the relation between ε_{eff} and N. Collisional kinetic theory concerns low density scalings, where the occupied volume $N{\varepsilon_{\text{eff}}}^3$ goes to 0. The good scaling limit looks to be the Boltzmann–Grad limit $N{\varepsilon_{\text{eff}}}^2 = 1$, at least for $s > 1$ (see [Gra49]). It corresponds to a regime such that on a time interval of length 1, any particle passes nearby another particle (at distance less than ε_{eff}, see Remark 1.1 for a comment on this scaling).

In the case $s > 1$, if at time 0 the particles are "sufficiently independent", the distribution of a typical particle $f(t, x, v)$ is a solution of the Boltzmann equation (introduced by Maxwell in 1867 [Max03] and Boltzmann in 1871 [Bol96])

$$\begin{gathered} \partial_t f + v \cdot \nabla_x f = Q_s(f, f) \\ Q_s(f,h)(v) := \int_{\mathbb{R}^3 \times \mathbb{S}^2} (f(v')h(v'_*) - f(v)h(v_*))\, b_s(v - v_*, \eta)\, \mathrm{d}v_*\, \mathrm{d}\eta, \\ v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2}\eta,\ v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2}\eta, \end{gathered} \tag{1.1}$$

where the kernel b_s depends on the potential $\mathscr{U}_s(\cdot)$. The collision operator Q can be interpreted as a jump operator for the velocities. The rigorous proof of such a theorem for long range interactions is now an open problem, and this paper can be interpreted as an intermediate result.

For a power law $\mathscr{U}_s$ with $s \geq 1$, the kernel b_s is equal to

$$b_s(z, \eta) = |z|^{\frac{s-4}{s}} q_s(z \cdot \eta), \text{ with } q_s(\cos\theta) \underset{\theta \sim 0}{\sim} K\theta^{-\frac{2+s}{s}} \tag{1.2}$$

for some constant K. Hence, the collision kernel is not integrable near the singularity $\eta \cdot \frac{v - v_*}{|v - v_*|} \simeq 1$ (when the collisions are grazing). We say that the Boltzmann kernel has no cutoff. However, the Boltzmann operator Q_s can be defined (everywhere) for $s > 1$ if the functions f and h are differentiable (see [Gou97, Vil98]).

In the Coulombian case $s = 1$, the singularity near $\theta \sim 0$ is too large to define the collision operator, even for smooth functions. In 1936, Landau proposed in [Lan36] an operator that describes the collision between Coulombian particles. His analysis was based on cutting the long range interaction, responsible for the "collisions" with small deviation angle θ. His argument was that in a neutral Coulomb gas, the interactions between ions (big positively charged particles) are screened by electrons (small negatively charged particles). Hence, the particles interact *via* the effective potential $\frac{\chi(|x|/R)}{|x|}$ where $\chi \simeq \mathbb{1}_{[0,1]}$ is a cutting function, and R is called the Debye radius[1]. In the present paper we study the limit $R \to \infty$. ($R \gg 1$). Landau finally obtained the collision operator $c(\log R)Q_L(f, f)$ where c is a diffusion coefficient, $\log R$ is the *Coulomb logarithm* coming from the singularity, and Q_L is defined by

$$Q_L(f,h)(v) := 2\pi\nabla_v \cdot \int_{\mathbb{R}^3} \frac{|v - v_*|^2 \mathrm{Id} - (v - v_*)^{\otimes 2}}{|v - v_*|^3} \big(\nabla f(v)h(v_*) - f(v)\nabla h(v_*)\big)\, \mathrm{d}v_*. \tag{1.3}$$

The Landau equation is

$$\partial_t f + v \cdot \nabla_x f = cQ_L(f, f). \tag{1.4}$$

Remark 1.1. *It is not clear if, in the case $s \leq 2$, the Boltzmann equation is really the good description in the Boltzmann-Grad limit $N{\varepsilon_{\text{eff}}}^2 = 1$ of the particle system with interaction through the long range potential $\mathscr{U}_s(\frac{x}{\varepsilon_{\text{eff}}}) = \frac{{\varepsilon_{\text{eff}}}^s}{|x|^s}$.*

Indeed, it can be compared with the mean-field regime $N{\varepsilon_{\text{eff}}}^2 = {\varepsilon_{\text{eff}}}^{2-s} \ll 1$ (the interaction potential is $\frac{1}{N}\mathscr{U}_s(x)$), which is much more dilute. Then, the mean-field effects should dominate the collisional effects. One can expect that, because of some "screening effect" (similar to that in the Coulomb case), particles interact via an effective potential with faster decay at infinity.

For hard potentials $s > 2$, the Boltzmann-Grad scaling is dominant, and the Boltzmann equation should be the true asymptotic description of the system. A similar discussion has been performed in [SW25] *for the definition of the equilibrium measure for particles interacting through the potential $\mathscr{U}_s(\frac{\cdot}{\varepsilon_{\text{eff}}})$.*

[1] In the Physics literature the constant R is typically interpreted as a normalisation constant, to be estimated

In this paper we will work with microscopically truncated interactions and therefore avoid the discussion of screening properties. Our purpose, however, is to obtain uncutoff equations by removing the truncation.

The first rigorous derivation of the Boltzmann equation was performed by Lanford for hard-spheres [Lan75] and by King [Kin75] for regular potentials with compact support.

Our strategy for deriving the Boltzmann (or Landau) equation associated with the potential $\mathscr{U}_s$ is to split the problem into two steps. First, we consider a system of particles interacting *via* the screened potential

$$\mathscr{U}_{s,R}(\tfrac{x}{\varepsilon_{\rm eff}}) := \chi(\tfrac{|x|}{R\varepsilon_{\rm eff}})\tfrac{\varepsilon_{\rm eff}{}^s}{|x|^s},$$

where R is a cutoff radius (which will go to infinity), $\chi(r) : \mathbb{R}^+ \to [0,1]$ is a smooth, decreasing cutoff function:

$$\chi(0) = 1,\ \chi([1,\infty[) = \{0\},\ \chi' \leq 0.$$

Taking the Boltzmann-Grad limit $N \to \infty$, $N\varepsilon_{\rm eff}{}^2 = 1$ ($N\varepsilon_{\rm eff}{}^2 = (\log R)^{-1}$ in the Coulomb case $s = 1$), we recover the cut-off Boltzmann equation. Defining a "collision" between two particles as the moment when they get closer than the effective radius $\varepsilon_{\rm eff}$, the Boltzmann-Grad scaling is chosen such that a particle has on average one collision per unite of time.

In a second time, we take the grazing collision limit $R \to \infty$ to pass from the cut-off Boltzmann equation to the Boltzmann equation associated with $\mathscr{U}_s$ if $s > 1$ (respectively, the Landau equation if $s = 1$).

If we obtain some quantitative estimates for the two regimes, one can hope to take the limits $R \to \infty, N \to \infty$ simultaneously, assuming that R grows much slower than N (we will need R of order $O((\log\log N)^{1/6})$).

The big difficulty is that we need the validity of the cut-off Boltzmann equation on a large time interval of order $O(1)$ (to be compared to the validity time $O(R^{-2})$ obtained by King [Kin75] for interaction potentials supported in a ball of radius R, see the next section). In order to get a long-time result, we look at a linearized version of the system near the thermodynamic equilibrium (or Gibbs state). This equilibrium can be defined as the probability law with density

$$M^N_{\varepsilon_{\rm eff},R}(X_N, V_N) := \frac{1}{\mathscr{Z}^N_{\varepsilon_{\rm eff},R}} \exp\left(-\sum_{i=1}^N \frac{|v_i|^2}{2} - \sum_{1\leq i<j\leq N} \mathscr{U}_{s,R}\left(\frac{x_i - x_j}{\varepsilon_{\rm eff}}\right)\right) \tag{1.5}$$

at positions $X_N = (x_1, \cdots, x_N)$ and velocities $V_N = (v_1, \cdots, v_N)$. The term $\mathscr{Z}^N_{\varepsilon_{\rm eff},R}$ is a normalization constant such that

$$\int M^N_{\varepsilon_{\rm eff},R}\, \mathrm{d}X_N\, \mathrm{d}V_N = 1.$$

We want to understand the fluctuation field $\zeta^t_{\varepsilon_{\rm eff}}$ around the equilibrium: for a test function g, we define

$$\zeta^t_{\varepsilon_{\rm eff}}(g) := \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^N g(\mathbf{x}_i^{\varepsilon_{\rm eff}}(t), \mathbf{v}_i^{\varepsilon_{\rm eff}}(t)) - \mathbb{E}_{\varepsilon_{\rm eff}}\left[\frac{1}{N}\sum_{i=1}^N g(\mathbf{x}_i^{\varepsilon_{\rm eff}}(t), \mathbf{v}_i^{\varepsilon_{\rm eff}}(t))\right]\right). \tag{1.6}$$

In the previous equality, $(\mathbf{x}_i^{\varepsilon_{\rm eff}}(t), \mathbf{v}_i^{\varepsilon_{\rm eff}}(t))$ denotes the position and velocity of the i-th particle at time t, and the expectation is taken with respect to the Gibbs measure $M^N_{\varepsilon_{\rm eff},R}\, \mathrm{d}X_N\, \mathrm{d}V_N$. The fluctuation field $\zeta^t_{\varepsilon_{\rm eff}}$ has been studied by Bodineau *et al* in the hard sphere system. They have shown $\zeta^t_{\varepsilon_{\rm eff}}$ verifies a central limit theorem, and converges to a Gaussian field ζ^t_0 described by the linearized Boltzmann equation (see [BGSR17, BGSRS23b, BGSRS24, LB25]). Our study is focused on the description of the covariance for more general interaction potential.

One can now write a first informal version of the theorem proved in the present paper (a rigorous version is written in Theorem 3[2]).

[2]The vague Claims 1 and 2 are written for simplicity of the presentation in the *canonical* setting, *id est* the number of particles is a constant N that goes to infinity. In fact, we will rather work in the *grand canonical* setting, where the number of particles is a random variable $\mathscr{N}$, with $\mathbb{E}\mathscr{N}$ going to infinity. The system becomes less rigid, and it avoids the apparition of additional error terms. However, one can expect that the canonical and grand canonical systems behave asymptotically in the same way (although the proof is missing).

Claim 1. *Consider a system of N particles evolving with respect to Newton's laws, interacting through the pairwise potential $\mathscr{U}_{s,R}(\cdot/\varepsilon_{\text{eff}})$, with $s \geq 1$. At time zero, the particles are distributed according to the Gibbs equilibrium measure. Parameters $N, \varepsilon_{\text{eff}}, R$ tuned as*

$$N \to \infty,\ R \to \infty,\ R = o\left((\log\log N)^{1/6}\right),\ \text{and } N{\varepsilon_{\text{eff}}}^2 = \begin{cases} (\log R)^{-1} & \text{if } s = 1, \\ 1 & \text{if } s > 1. \end{cases}$$

Fix g and h two test functions. Then denoting $M(v) := \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{3/2}}$, we have that

$$\mathbb{E}_{\varepsilon_{\text{eff}}}\left[\zeta^t_{\varepsilon_{\text{eff}}}(g)\zeta^0_{\varepsilon_{\text{eff}}}(h)\right] \underset{\varepsilon\to 0}{\longrightarrow} \int \mathbf{g}(t,x,v)h(x,v)M(v)\,\mathrm{d}x\,\mathrm{d}v \tag{1.7}$$

with $\mathbf{g}(t,x,v)$ the solution of the linearized equation

$$\begin{cases} \partial_t \mathbf{g} + v\cdot\nabla_v \mathbf{g} = \mathscr{L}_\infty \mathbf{g}, \\ \mathbf{g}(t=0,x,v) = g(x,v) \end{cases} \text{where } \mathscr{L}_\infty \mathbf{g} := \begin{cases} \frac{1}{M}\left(Q_L(Mg,M) + Q_L(M,Mg)\right) \text{ if } s = 1, \\ \frac{1}{M}\left(Q_s(Mg,M) + Q_s(M,Mg)\right) \text{ if } s > 1. \end{cases} \tag{1.8}$$

1.1. **Modification of the scaling parameters and the sub-Coulomb case ($s \in [0,1)$).** For a fix $s \geq 1$ and a cut-off function $\chi : \mathbb{R}^+ \to [0,1]$, we define the *interaction radius* $\varepsilon := R\varepsilon_{\text{eff}}$, the *coupling constant* $\alpha = R^{-s}$ and the interaction potential

$$\mathscr{V}(x) := \frac{\chi(|x|)}{|x|^s}.$$

Hence, we have the equality $\mathscr{U}_{s,R}(x/\varepsilon_{\text{eff}}) = \alpha\mathscr{V}(x/\varepsilon)$.

The scaling parameters (ε,α) and $(\varepsilon_{\text{eff}}, R)$ are two different parameterizations of the same system, and taking

$$N \to \infty,\ R \to \infty,\ R = o\left((\log\log N)^{1/6}\right),\ \text{and } N{\varepsilon_{\text{eff}}}^2 = \begin{cases} (\log R)^{-1} & \text{if } s = 1 \\ 1 & \text{if } s > 1 \end{cases}$$

is equivalent to taking

$$N \to \infty,\ \alpha \to 0,\ \frac{1}{\alpha} = o\left((\log\log N)^{s/12}\right),\ \text{and } N\varepsilon^2 = \begin{cases} \alpha^{-2}|\log\alpha|^{-1} & \text{if } s = 1 \\ \alpha^{-2/s} & \text{if } s > 1 \end{cases}$$

In the core of the proof (from Section 2 to the end), we prefer using (ε,α) as it will simplify the notation. However, the couple $(\varepsilon_{\text{eff}}, R)$ may be more natural to describe a gas interacting *via* a power law $\mathscr{U}_s$, $s \geq 1$.

Another advantage of taking the parameters (ε,α) is that it allows a natural generalization of the result to potentials in the sub-Coulomb range $s \in [0,1)$ which is the microscopic interpretation of the so-called *grazing collision limit.*

We look at a system of N particles interacting pairwise *via* the potential $\alpha\mathscr{V}(\cdot/\varepsilon)$. At time $t = 0$, they are distributed with respect to the canonical Gibbs measure

$$M^N_{\varepsilon,\alpha}(X_N, V_N) := \frac{1}{\mathscr{Z}_{N,\varepsilon,\alpha}} \exp\left(-\sum_{i=1}^N \frac{|v_i|^2}{2} - \alpha \sum_{1\leq i<j\leq N} \mathscr{V}\left(\frac{x_i - x_j}{\varepsilon}\right)\right),$$

where $\mathscr{Z}_{N,\varepsilon,\alpha}$ is a normalization constant, and ζ_ε the fluctuation field (defined in the same way as $\zeta_{\varepsilon_{\text{eff}}}$). Then we can write an informal theorem (with the rigorous version written in Theorem 3):

Claim 2. *Consider a system of N particles evolving according to Newton's laws, interacting through the pairwise potential $\alpha\mathscr{V}(\cdot/\varepsilon)$, where $s \in [0,1)$. At time zero, the particles are distributed with respect to the equilibrium measure. Parameters N, ε, α are tuned as*

$$N \to \infty,\ \alpha \to 0,\ \tfrac{1}{\alpha} = o((\log\log N)^{1/12}),\ \text{and } N\varepsilon^2 = \alpha^{-2}.$$

Fix g and h two test functions. Then, denoting $M(v) := \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{3/2}}$,

$$\mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(g)\zeta^0_\varepsilon(h)\right] \underset{\varepsilon\to 0}{\longrightarrow} \int \mathbf{g}(t,x,v)h(x,v)M(v)\,\mathrm{d}x\,\mathrm{d}v \tag{1.9}$$

with $\mathbf{g}(t,x,v)$ *the solution of the linearized equation*

$$\begin{cases} \partial_t \mathbf{g} + v\cdot\nabla_v \mathbf{g} = \frac{c_{\mathscr{V}}}{M}\Big(Q_L(Mg,M)+Q_L(M,Mg)\Big), \\ \mathbf{g}(t=0,x,v) = g(x,v), \end{cases}$$

where $c_{\mathscr{V}}$ *is a diffusion constant defined by*

$$c_{\mathscr{V}} = \frac{1}{16\pi^2}\int_{\mathbb{R}^3} \delta(k\cdot\vec{e}_1)|k|^2|\hat{\mathscr{V}}(k)|^2\,\mathrm{d}k \tag{1.10}$$

where $\vec{e}_1\in\mathbb{S}^2$ *an unitary vector and* $\hat{\mathscr{V}}(k)$ *is the Fourier transform of* $\mathscr{V}$ *using the convention*

$$\hat{\mathscr{V}}(k) = \int \mathscr{V}(x)e^{-ik\cdot x}\,\mathrm{d}x.$$

Remark 1.2. *The Landau equation is usually derived from interacting particle systems in the* weak coupling limit*: one fixes*

$$\alpha = \varepsilon^{1/2}, \quad N\varepsilon^3 = 1.$$

For example, Bobylev, Pulvirenti, and Saffirio provided in [BPS13] *a consistency result (a result at time* 0*) for smooth interaction in this scaling.*

In the present article, we are only able to treat the cases

$$\alpha^{-1} = o(\log|\log\varepsilon|)^{1/12} \ll \varepsilon^{-1/2}, \quad N\varepsilon^2\alpha^2 = 1,$$

which are far from the weak coupling limit. The same Landau equation is, however, expected to hold in all intermediate regimes connecting Boltzmann-Grad and weak-coupling scalings (see on that subject [NSV18, NVW22, NVW21, PS21]*).*

1.2. **State of the art.** Now we recall some results about the derivation of the Boltzmann and Landau equation from a particle system .

In the nonlinear setting, the only results hold for potentials $\mathscr{U}(\cdot)$ supported in a ball $\{x\in\mathbb{R}^3,\ |x|\le R\}$. In the Boltzmann-Grad scaling $N{\varepsilon_{\text{eff}}}^2 = 1$, the distribution of a typical particle follows the Boltzmann equation up to a time $O(1/R^2)$. The first derivation was performed by Lanford [Lan75] for hard spheres (*i.e.* $\exp(-\mathscr{U}_{\text{hs}}(r)) = \mathbb{1}_{r>1}$) and King [Kin75] for more general compactly supported potentials (see also [GSRT13, PSS14, Den18, BGSRS18]). The previous results have two defects. They are valid only up to a small time (for the atmosphere at temperature 20°C and pressure 10^5Pa, the validity time is $10^{-9}s$), and the results apply only to a compactly supported interaction potential. A first long-time result out of equilibrium is [IP89], in a setting where the dispersive effects are dominant. More recently, Deng, Hani and Ma provide a long time result out of equilibrium [DHM24] for more general initial datum f_0. Their result holds up to the minimum between the existing time of $f(t)$, the solution of the Boltzmann equation with initial data f_0, and a time $O(\log|\log N|^\alpha)$ for some $\alpha\in(0,1)$ given by a sampling time strategy.

For the Landau equation, the unique results hold only at time 0 (see [BPS13, Win21, Due21]): the authors obtain the equality

$$(\partial_t f)_{|t=0} = -v\cdot\nabla_x f_0 + Q_L(f_0,f_0).$$

Note that in [BPS13, Win21], the authors do not look at the real particle system but at a simplified version (they truncate the BBGKY hierarchy). In [Due21], Duerinckx proves the consistency of the Lenard–Balescue equation, which can be understood as a modification of the Landau equation. However, the scaling is far from the collisional scaling that we are treating here.

This is not the first attempt to derive a linear version of the Boltzmann or Landau equations.

A linear equation can be obtained from the study of the *Lorentz gas*: one fixes a background of obstacles distributed with respect to the Poisson measure of parameter N. One follows a unique particle that interacts pairwise with the obstacles *via* the potential $\mathscr{U}(\cdot/\varepsilon_{\text{eff}})$ with $N\varepsilon^2 = 1$. In the limit $N\to\infty$, the density of the tagged particle $f(t,x,v)$ follows the linear Boltzmann equation

$$\partial_t f(t,x,v) + v\cdot\nabla_x f(t,x,v) = \int_{\mathbb{S}^2} \left(f(t,x,|v|\eta) - f(t,x,v)\right) b(\eta\cdot\tfrac{v}{|v|})\,\mathrm{d}\eta \tag{1.11}$$

where b is a collision kernel. It does not depend on the norm of v as the obstacles are fixed. This system was first described by Gallavotti [Gal69] in the case of the hard spheres, and later adapted for the derivation of the linear Boltzmann equation without cutoff [DP99] and linear Landau equation [DR01]. Note that in [DP99], the authors used, in the same way as this paper, interaction through a screened potential $\tilde{\mathscr{U}}(x) = \mathbb{1}_{|x|\le R}\left(|x|^{-s} - R^{-s}\right)$ for $s>1$.

A second possibility is the treatment of the *linear particle setting*. We look at a system of N interacting particles, initially distributed with respect to the Gibbs measure (1.5). Then, one wants to follow a tagged particle of the system, id est compute the covariance (for g, h two test functions)

$$\lim_{N\to\infty} \mathbb{E}_{\varepsilon_{\rm eff}}\left[h(\mathbf{x}_1^{\varepsilon_{\rm eff}}(t),\mathbf{v}_1^{\varepsilon_{\rm eff}}(t))g(\mathbf{x}_1^{\varepsilon_{\rm eff}}(0),\mathbf{v}_1^{\varepsilon_{\rm eff}}(0))\right] = \int h(x,v)\mathbf{g}(t,x,v)M(v)\,\mathrm{d}v\,\mathrm{d}x \tag{1.12}$$

where $\mathbf{g}(t,x,v)$ is solution of the linear Boltzmann equation

$$\partial_t g(t,x,v) + v\cdot\nabla_x g(t,x,v) = \frac{1}{M(v)} Q_{\mathscr{U}}(Mg,M), \tag{1.13}$$

where $Q_{\mathscr{U}}$ is the Boltzmann operator associated with the potential $\mathscr{U}$. The first long-time result on such a system was obtained in [vBLLS80]. Later, Bodineau *et al.* [BGSR17] provided a quantitative proof in the hard sphere setting. This proof has been adapted to other potentials in order to derive the linear Landau equation [Cat18] or the linear Boltzmann equation without cutoff [Ayi17]. In her paper, Ayi does not consider interactions through the cut-off potential $\mathscr{U}_{s,R}$, but directly a long-range potential $\mathscr{U}$ with fast decay at infinity (she needs $\mathscr{U}(r)\le O(\exp(-\exp\exp|x|^4))$). Up to our knowledge, it is the unique result where the particles interact through genuine infinite potential.

While the linear setting is a $O(1)$ perturbation of equilibrium, the linearized setting (which is treated in the present paper) is a $O(N)$ perturbation of equilibrium. In [Spo81, Spo83], Spohn showed that the fluctuation field $\zeta_{\varepsilon_{\rm eff}}$ formally verifies a Central Limit Theorem in the limit $N\to\infty$. The first step to prove such a theorem is to compute the covariance of the fluctuation field. This has been performed in the hard sphere setting by Bodineau *et al.* in [BGSR17] in dimension 2 and [BGSRS23b] in any dimension bigger than 3 (note also [LB25]). Then, they proved that the limit is Gaussian in [BGSRS24].

2. Definition of the system and strategy of the proof

2.1. The Hamiltonian dynamics.

Let $\mathbb{T} := \mathbb{R}^d/\mathbb{Z}^d$ (with $d\ge 2$) be the domain. We denote $\mathbb{D} = \mathbb{T}\times\mathbb{R}^d$ its tangent bundle and $\mathbb{D}^n$ the n-particle canonical phase space. In the following, we use the notation

$$X_n = (x_1,\cdots,x_n)\in\mathbb{T}^n,\ V_n = (v_1,\cdots,v_n)\in\mathbb{R}^{nd},\ \text{and } z_i = (x_i,v_i)\in\mathbb{D}.$$

On each $\mathbb{D}^n$, we construct the Hamiltonian dynamics associated with the Energy

$$\mathscr{H}_n(Z_n) := \frac{1}{2}|V_n|^2 + \mathscr{V}_n(X_n),\quad \mathscr{V}_n(X_n) := \sum_{1\le i<j\le n} \alpha\mathscr{V}\left(\frac{|x_i-x_j|}{\varepsilon}\right), \tag{2.1}$$

$$\forall i\in[1,n],\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}x_i = \nabla_{v_i}\mathscr{H}_n(Z_n(t)) = v_i,\\ \frac{\mathrm{d}}{\mathrm{d}t}v_i = -\nabla_{x_i}\mathscr{H}_n(Z_n(t)) = \frac{\alpha}{\varepsilon}\sum_{\substack{j=1\\ j\neq i}}^n \nabla\mathscr{V}\left(\frac{x_i-x_j}{\varepsilon}\right).\end{cases} \tag{2.2}$$

We impose the following condition on the interaction potential

Assumption 2.1. *There exists a constant $s\in[0,\infty)$ and a decreasing cut-off function $\chi\in\mathscr{C}([0,\infty))\cap\mathscr{C}^2([0,1))$ such that*

$$\mathscr{V}(x) := \frac{\chi(|x|)}{|x|^s},\ \chi(0)=1,\ \chi([1,\infty)) = \{0\}. \tag{2.3}$$

This dynamics is well defined for all times, almost everywhere in $\mathbb{D}^n$ with respect to the Lebesgue measure.

2.2. Grand-canonical ensemble and stationary measure.

In the following, we choose not to fix the number of particles $\mathscr{N}$, and suppose that $\mathscr{N}$ is a random variable with $\mathbb{E}_\varepsilon[\mathscr{N}]$ going to infinity when ε goes to 0 (we say that we consider the *grand canonical ensemble*). It will simplify a lot of computations[3].

We denote $\mathscr{D} := \bigsqcup_{n\ge 0}\mathbb{D}^n$ the grand canonical phase space. We can then extend the Hamiltonian dynamics to $\mathscr{D}$ and denote $\mathbf{Z}_{\mathscr{N}}(t)$ the realization (defined almost surely) of the Hamiltonian flow on $\mathscr{D}$ with random initial data $\mathbf{Z}_{\mathscr{N}}(0)$: for $\mathscr{N}=n$, $\mathbf{Z}_{\mathscr{N}}(t)$ follows the Hamiltonian dynamics on $\mathbb{D}^n$.

[3]The treatment of the canonical setting remains an open problem.

The initial data are sampled according to the stationary measure introduced now. The *grand canonical Gibbs measure* $\mathbb{P}_\varepsilon$ (and its expectation $\mathbb{E}_\varepsilon$) are defined on $\mathscr{D}$ as follows: an application $H : \mathscr{D} \to \mathbb{R}$ is a test function if there exists a sequence $(h_n)_{n\geq 0}$ with $h_n \in L^\infty(\mathbb{D}^n)$ and

$$\text{if } \mathscr{N} = n,\ \mathbf{Z}_{\mathscr{N}} = (\mathbf{z}_1, \cdots, \mathbf{z}_n),\ H(\mathbf{Z}_{\mathscr{N}}) := h_n(\mathbf{z}_1, \cdots, \mathbf{z}_n).$$

Fixing $\mu > 0$ the *chemical potential*[4], we define $\mathbb{E}_\varepsilon$ as

$$\mathbb{E}_\varepsilon[H(\mathbf{Z}_{\mathscr{N}})] := \frac{1}{\mathscr{Z}} \sum_{n\geq 0} \frac{\mu^n}{n!} \int_{\mathbb{D}^n} h_n(Z_n)\, \frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{nd/2}}\, \mathrm{d}Z_n, \tag{2.4}$$

where $\mathscr{Z}$ is a normalisation constant such that $\mathbb{E}_\varepsilon[1] = 1$. The *mean free path* $\mathfrak{d}$ is defined by

$$\mathfrak{d} = \frac{1}{\mu\, \varepsilon^{d-1}}. \tag{2.5}$$

It can be interpreted as the typical distance crossed by a particle between two collisions.

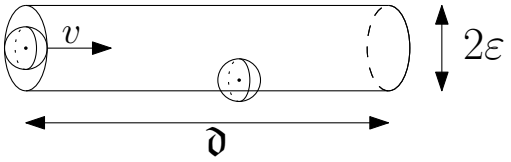

FIGURE 1. The first particle will meet the second one. Here v is of order 1.

We will consider only ε small and $\mathfrak{d} \underset{\varepsilon\to 0}{=} O(1)$, for which the series defining the Gibbs measure converges absolutely.

The empirical distribution at time t is defined as the average configuration of particles at time t: for any g test function on $\mathbb{D}$,

$$\pi_t^\varepsilon(g) := \frac{1}{\mu} \sum_{i=1}^{\mathscr{N}} g(\mathbf{z}_i(t)). \tag{2.6}$$

At equilibrium, we have the following law of large numbers. Denote

$$M(v) := \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{\frac{d}{2}}}. \tag{2.7}$$

Proposition 2.2. *For any continuous and bounded test function $g : \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}$, for all $t \in \mathbb{R}$ and for any $\delta > 0$,*

$$\lim_{\varepsilon\to 0} \mathbb{P}_\varepsilon \left[\left| \pi_\varepsilon^t(g) - \int g(z) M(v)\, \mathrm{d}z \right| \geq \delta \right] = 0. \tag{2.8}$$

Remark 2.1. *The previous result is a simple corollary of the Lanford theorem and of the stationarity in time of the measure (see* [Lan75, Kin75]*). From Proposition 4.1, we can deduce the L^2 counterpart of this law of large numbers.*

The aim of this article is to investigate the next order, namely the *fluctuation field*[5]

$$\zeta_\varepsilon^t(g) := \mu^{\frac{1}{2}} \Big(\pi_\varepsilon^t(g) - \mathbb{E}_\varepsilon[\pi_0^\varepsilon(g)] \Big). \tag{2.9}$$

2.3. **Binary collision, scattering, and definition of the linearized Boltzmann operator.** Interactions involving more than two particles become negligible in the Boltzmann-Grad limit.

The present section is dedicated to describing the map between pre-collisional and post-collisional velocities. It is called the *scattering map* (see Chapter 8 of [GSRT13] for a more detailed discussion).

Consider two interacting particles 1 and 2 following the Hamiltonian dynamic associated with $\mathscr{H}_2$. At time 0, particles have coordinates $(X_2(0), V_2(0))$ with

$$x_1(0) = \varepsilon\nu,\ x_2(0) = 0,\ v_1 = v \text{ and } v_2(0) = v_*$$

where $\nu \in \mathbb{S}^{d-1}$ and $(v - v_*) \cdot \nu > 0$.

[4]It takes the role of the number of particle N has in the canonical setting.

[5]As the Gibbs measure is stationary in time, for all $t \in \mathbb{R}$, $\mathbb{E}_\varepsilon[\pi_0^\varepsilon(g)] = \mathbb{E}_\varepsilon[\pi_\varepsilon^t(g)]$.

The particles will interact on a finite interval $[0,[\tau]]$ with $[\tau]$ the infimum of $\{\tau>0,\ |x_2(\tau)-x_1(\tau)|>\varepsilon\}$. The time $[\tau]$ is finite and bounded by $\frac{\varepsilon|v_1-v_2|}{|\nu\times(v_1-v_2)|^2}$, with $\times$ the cross product (see Lemma B.1). We define (ν',v',v'_*) as

$$\nu':=\frac{x_2(\tau)-x_1(\tau)}{\varepsilon}\text{ and }(v',v'_*):=\big(v_1([\tau]),v_2([\tau])\big).$$

In addition, the scattering conserves both momentum, kinetic energy, and angular momentum:

$$v+v_*=v'+v'_*,\ |v|^2+|v_*|^2=|v'|^2+|v'_*|^2\text{ and }(v-v_*)\times\nu=(v'-v'_*)\times\nu' \tag{2.10}$$

We deduce that

$$|(v-v_*)\cdot\nu|=|(v'-v'_*)\cdot\nu'|. \tag{2.11}$$

Definition 2.1. *The* scattering application *defined by*

$$\xi_\alpha:(\nu,v,v_*)\mapsto(\nu',v',v'_*) \tag{2.12}$$

is a local diffeomorphism which sends the measure $\mathrm{d}v\mathrm{d}v_*\mathrm{d}\nu$ *to* $\mathrm{d}v'\mathrm{d}v'_*\mathrm{d}\nu'$. *In addition, it does not depend on the particle radius* ε.

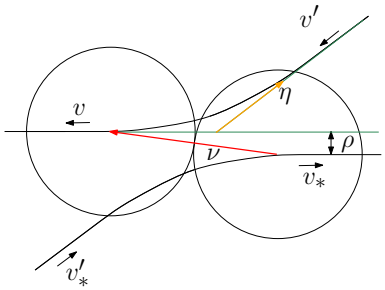

FIGURE 2. The scattering between two particles.

We define the linearized Bolzmann operator in the King's form:

$$\mathscr{L}_{\mathscr{U}}g(v):=\int_{\mathbb{S}\times\mathbb{R}^d}\big(g(v')+g(v'_*)-g(v)-g(v_*)\big)((v-v_*)\cdot\nu)_+M(v_*)\,\mathrm{d}\nu\,\mathrm{d}v_*, \tag{2.13}$$

where we apply the scattering with interaction potential $\mathscr{U}(\cdot)$, and $\mathscr{L}_\alpha:=\mathscr{L}_{\alpha\mathscr{V}}$.

This operator describes the variation of mass in a gas due to changes of velocity of colliding particles. It is well known that the operator $\mathscr{L}_\alpha$ is a self-adjoint non-positive operator on $L^2(M(v)\,\mathrm{d}z)$.

Remark 2.2. *We say that the Boltzmann operator* $\mathscr{L}_{\mathscr{U}}$ *has a cutoff because we truncate the long range interaction.*

There is another interpretation of this property. For parameters (v,v_*,ν), *we can define the vector* η *such that*

$$v'=\frac{v+v_*}{2}+\frac{|v-v_*|}{2}\eta,\ v'_*=\frac{v+v_*}{2}-\frac{|v-v_*|}{2}\eta,$$

and $b_\alpha(v-v_*,\eta)$ *(called the collision kernel) the Jacobian of the application* $\nu\mapsto\eta$*:*

$$((v-v_*)\cdot\nu)_+\,\mathrm{d}\nu\to b_\alpha(v-v_*,\eta)\,\mathrm{d}\eta.$$

We say that the Boltzmann operator has a cutoff because for any $v-v_*$, *the following bound holds*

$$\int b_\alpha(v-v_*,\eta)\,\mathrm{d}\eta<\infty.$$

See Chapter 8 of [GSRT13], *Appendix of* [PSS14] *and Proposition 2.3.2 of* [LB23] *for explicit estimations of the collision kernel* b_α *for certain class of interaction potential.*

2.4. **Convergence to the linearized Boltzmann equation with a cut-off.** We recall that we have divided the proof of Claim 1 into two steps. The first step is the Boltzmann-Grad limit $\mu \to \infty$. As we want to take the *grazing collision* limit $\alpha \to 0$ in a second step, we need a quantitative rate of convergence.

We define the norm

$$\|g\|_0 := \sup_{(x,v)\in\mathbb{D}} \left|M^{-1}(v)g(x,v)\right| \text{ and } \|g\|_k := \sum_{|\alpha|\le k} \|\nabla^\alpha g\|_0. \tag{2.14}$$

Theorem 1. *Let g and h be two test functions $\mathscr{C}^1(\mathbb{D})$, with $\|g\|_1, \|h\|_1 < \infty$. Then there exist three constants $C > 1$, $C' > 1$ and $\mathfrak{a} \in (0,1)$ depending only on the dimension such that for any ε small enough, $T > 1$, $\theta < \frac{1}{C'T^2}$, $\alpha, \mathfrak{d} \in (\log|\log\varepsilon|^{-1}, 1)$,*

$$\sup_{t\in[0,T]} \left| \mathbb{E}_\varepsilon\left[\zeta_\varepsilon^t(h)\zeta_\varepsilon^0(g)\right] - \int h(z)\mathbf{g}_\alpha(t,z)M(z)\,\mathrm{d}z \right| \le C\left(\frac{\theta T^2}{\mathfrak{d}^3} + \varepsilon^{\mathfrak{a}}\left(\frac{CT}{\mathfrak{d}}\right)^{2^{\frac{T}{\theta}+1}}\right)\|g\|_1\|h\|_1, \tag{2.15}$$

where $\mathbf{g}_\alpha(t,z)$ is the solution in $L^\infty(\,\mathrm{d}t, L^2(M(v)\,\mathrm{d}z))$ of the linearized Boltzmann equation

$$\begin{aligned} \partial_t \mathbf{g}_\alpha(t) + v\cdot\nabla_x \mathbf{g}_\alpha(t) &= \frac{1}{\mathfrak{d}}\mathscr{L}_\alpha \mathbf{g}_\alpha(t), \\ \mathbf{g}_\alpha(t=0) &= g \end{aligned} \tag{2.16}$$

Remark 2.3. *It is classical that there is a unique solution to the linearized Boltzmann equation, which is bounded globally in time in $L^2(M(v)\,\mathrm{d}z)$ (see e.g., Section 7 in* [CIP94]*).*

The theorem is valid in any dimension $d \ge 2$. Its proof is the main purpose of the present article. We conclude the proof of Theorem 1 by Estimation (3.26), and we outlined the main step of the proof in Section 2.6.

2.5. **Derivation of the linearized Landau equation and Boltzmann equation without cut-off.** In this section, we discuss the second step, namely the grazing collision limit. We fix the dimension at $d = 3$ as it is the physical case. We only state the main results, as the proof can be found in the joint article [LBW22].

The case where the singularity $\frac{1}{r^s}$ of the potential is stronger than the Coulomb singularity ($s > 1$). In the limit $\alpha \to 0$, we will only see the effects of the singularity at the origin. We define the power law potential $\mathscr{U}_s(r) := 1/r^s$. It is natural to guess that one has convergence of the Boltzmann operators

$$\alpha^{-\frac{2}{s}}\mathscr{L}_\alpha \to \mathscr{L}_{\mathscr{U}_s} \tag{2.17}$$

which is a linearized Boltzmann operator without cutoff (see Appendix A for a rigorous definition of $\mathscr{L}_{\mathscr{U}_s}$ and a justification of the scaling $\mathfrak{d} = \alpha^{2/s}$).

Remark 2.4. *We say that the Boltzmann operator $\mathscr{L}_{\mathscr{U}_s}$ has no cutoff because particles can interact at long range, and the collision kernel $b_s(v - v_*, \eta)$ associated to the potential $1/r^s$ (defined in Remark 2.2) is not integrable in the η variable (see* (1.2)*).*

The Coulomb case $s = 1$. It is not possible to define the Boltzmann operator for the Coulomb potential. However, we can prove (see [LBW22]) that for g a test function smooth enough,

$$\frac{1}{\alpha^2|\log\alpha|}\mathscr{L}_\alpha g \underset{\alpha\to 0}{\longrightarrow} c_{\mathscr{V}}\mathscr{K}g \tag{2.18}$$

where $c_{\mathscr{V}} = 1$ is a diffusion constant and $\mathscr{K}$ is the linearized Landau operator

$$\mathscr{K}g(v) = \frac{2\pi}{M(v)}\nabla_v\cdot\left(\int_{\mathbb{R}^3}\frac{P^\perp_{v-v_*}}{|v-v_*|}(\nabla g(v) - \nabla g(v_*))M(v)M(v_*)\,\mathrm{d}v_*\right). \tag{2.19}$$

Treat now the weak singularity $s \in [0,1)$. For these potentials, the scaling and the diffusion constant change:

$$\mathfrak{d} = \alpha^2, \quad 2\pi c_{\mathscr{V}} = \frac{1}{8\pi}\int_{\mathbb{R}^3}\delta(k\cdot\vec{e}_1)|k|^2|\hat{\mathscr{V}}(k)|^2\,\mathrm{d}k, \tag{2.20}$$

where $\vec{e}_1$ is a unit vector, and we use the convention $\hat{\mathscr{V}}(k) = \int_{\mathbb{R}^3} e^{-ik\cdot x}\mathscr{V}(x)\,\mathrm{d}x$ for the Fourier transform of $\mathscr{V}$. Then

$$\frac{1}{\alpha^2}\mathscr{L}_\alpha g \underset{\alpha\to 0}{\longrightarrow} c_{\mathscr{V}}\mathscr{K}g \tag{2.21}$$

The previous discussion can be summarized by the following theorem:

Theorem 2 (L.B.-Winter, [LBW22]). *For $g : \mathbb{D} \to \mathbb{R}$ smooth and $\mathscr{V}$ respecting Assumption 2.1, there exists a positive constant C such that*

$$\left\|\mathfrak{d}_{s,\alpha}^{-1}\mathscr{L}_\alpha g - \mathscr{L}_\infty g\right\|_{L^2(M(v)\,\mathrm{d}z)} \leq \frac{C}{|\log\alpha|}\|g\|_3, \tag{2.22}$$

where $\mathscr{L}_\infty$ and $\mathfrak{d}_{s,\alpha}$ are given by

Singularity	*Mean free-path*	*Limiting operator*
$s > 1$	$\mathfrak{d}_{s,\alpha} := \alpha^{2/s}$	$\mathscr{L}_\infty = \mathscr{L}_{\mathscr{U}_s}$
$s = 1$	$\mathfrak{d}_{s,\alpha} := \alpha^2\lvert\log\alpha\rvert$	$\mathscr{L}_\infty = \mathscr{K}$
$0 \leq s < 1$	$\mathfrak{d}_{s,\alpha} := \alpha^2$	$\mathscr{L}_\infty = c_{\mathscr{V}}\mathscr{K}$, $c_{\mathscr{V}} = \frac{1}{16\pi^2}\int_{\mathbb{R}^3}\delta(k\cdot\vec{e}_1)\lvert k\rvert^2\lvert\hat{\mathscr{V}}(k)\rvert^2\,\mathrm{d}k$

In addition, defining $\mathbf{g}_\infty(t)$ the solution of

$$\begin{aligned}\partial_t\mathbf{g}_\infty(t) + v\cdot\nabla_x\mathbf{g}_\infty(t) &= \mathscr{L}_\infty\mathbf{g}_\infty(t),\\ \mathbf{g}_\infty(t=0) &= g\end{aligned} \tag{2.23}$$

and $\mathbf{g}_\alpha$ the solution of (2.16) with $\mathfrak{d} := \mathfrak{d}_{s,\alpha}$, the following convergence holds

$$\mathbf{g}_\alpha \underset{\alpha\to 0}{\overset{*}{\rightharpoonup}} \mathbf{g}_\infty \text{ in } L^\infty_t(\mathbb{R}^+(L^2(M(v)\,\mathrm{d}z)). \tag{2.24}$$

Combining it with Theorem 1 we obtain the main theorem:

Theorem 3. *Let $f, g \in L^2(M(v)\,\mathrm{d}z)$ be two test functions.*
Consider a potential $\mathscr{V}$ such that the Assumptions 2.1 are verified and $\mathscr{V}(r) \underset{r\to 0^+}{\sim} \frac{1}{r^s}$, $s \geq 0$.
Fix the scaling $\mu\varepsilon^2\mathfrak{d}_{s,\alpha} = 1$. Then we have the following convergence result: for all $t \geq 0$,

$$\mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(h)\zeta^0_\varepsilon(g)\right] \underset{\substack{\alpha\to 0\\ \alpha\log|\log\varepsilon|^{\frac{\tilde{s}}{6}}\to\infty}}{\overset{\varepsilon\to 0}{\longrightarrow}} \int h(z)\mathbf{g}_\infty(t,z)M(z)\,\mathrm{d}z$$

where $\tilde{s} = \max(s,1)$ and $\mathbf{g}_\infty(t)$ is the solution of the equation (2.23).

Remark 2.5. *We recall that in the sup-Coulomb case ($s > 1$), taking the limit*

$$\varepsilon \to 0,\ \alpha \to 0,\ \mu\,\varepsilon^2 = \mathfrak{d}^{-1}_{s,\alpha} = \alpha^{-2/s} \text{ and } \alpha\log|\log\varepsilon|^{\frac{s}{6}} \to \infty$$

is equivalent to taking the limit

$$\varepsilon_{\mathrm{eff}} \to 0,\ R \to \infty,\ \mu\,{\varepsilon_{\mathrm{eff}}}^2 = 1 \text{ and } \frac{\log|\log\varepsilon|^{\frac{1}{6}}}{R} \to \infty.$$

Proof of Theorem 3. First, the space $E := \{g : \mathbb{D} \to \mathbb{R},\ \|g\|_1 < \infty\}$ is dense in $L^2(M(v\,\mathrm{d}z))$.
Since the two bilinear operators

$$(h,g) \mapsto \mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(h)\zeta^0_\varepsilon(g)\right],\ (h,g) \mapsto \int h(z)\mathbf{g}_\infty(t,z)M(z)\,\mathrm{d}z$$

are continuous on $L^2(M(v)\,\mathrm{d}z)$ (see [BGSRS23b]), it is sufficient to take $g, h \in E$.
Set $T := \max(1,t)$. Fixing $\theta := \frac{1}{\beta\log|\log\varepsilon|}$ for $\beta \in (0,1)$ small enough,

$$C\left(C\frac{T^2\theta}{\mathfrak{d}^3_{s,\alpha}} + \frac{T}{\theta}2^{T^2/\theta^2}\left(\frac{CT}{\mathfrak{d}_{s,\alpha}}\right)^{2^{T/\theta}}\varepsilon^{\mathfrak{a}/2}\right) = o(1)$$

Hence Theorem 1 provides

$$\mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(g)\zeta^0_\varepsilon(h)\right] = \int h(z)\mathbf{g}_\alpha(t,z)M(v)\,\mathrm{d}z + o(\|g\|_1\|h\|_1).$$

Theorem 2 provides the convergence

$$\int h(z)\mathbf{g}_\alpha(t,z)M(v)\,\mathrm{d}z \to \int h(z)\mathbf{g}_\infty(t,z)M(v)\,\mathrm{d}z.$$

This concludes the proof. □

Notations. For $\omega \subset \mathbb{N}$ a finite subset and $r \leq |\omega|$, we denote $\mathscr{P}_\omega^r$ the set of the unordered partitions $(\rho_1, \cdots, \rho_r)$ of the set ω.

For $m < n$ two integers, we denote $[m,n] := \{m, m+1, \cdots, n\}$ and $[n] := [1,n]$.

For $Z_n \in \mathbb{D}^n$, and $\omega \subset [n]$, we denote

$$Z_\omega := (z_{\omega(1)}, \cdots, z_{\omega(|\omega|)})$$

where $\omega(i)$ is the i-th element of ω counted in increasing order.

Given a family of particles indices $\{i_1, \cdots, i_n\}$, the notation $(i_1, \cdots, i_n)$ indicates the ordered sequence in which $\forall k \neq l$, $i_k \neq i_l$. In addition,

- $\underline{i}_n := (i_1, \cdots, i_n)$,
- for $m \leq n$, $\underline{i}_m = (i_1, \cdots, i_m)$, and more generally for $\omega \subset [1,n]$, $\underline{i}_\omega := (i_{\min \omega}, \cdots, i_{\max \omega})$,
- for $0 \leq m < n$ and $(i_1, \cdots, i_m)$, $\sum_{(i_{m+1}, \cdots, i_n)}$ denotes the sum over every family $(i_{m+1}, \cdots, i_n)$ such that for $1 \leq k < l \leq n$, $i_k \neq i_l$, and
$$\sum_{\underline{i}_n} = \sum_{(i_1, \cdots, i_n)},$$
- $\mathbf{Z}_{\underline{i}_n} := (\mathbf{z}_{i_1}, \cdots, \mathbf{z}_{i_n})$, as an ordered sequence.

We also precise the Landau[6] notation: $A = B + O(D)$ means that there exists a constant C depending only on the dimension such that $|A - B| < C\,D$. We denote $A \lesssim B$ if $A = O(B)$.

When we perform estimations, C is a positive constant depending only on the dimension (which can change from one line to another), and the final time t is supposed to be bigger than 1 (in general, we prefer to denote τ any intermediate time).

Finally, let h_n be a function on $\mathbb{D}^n$. We denote (by a slight abuse of notation)

$$\mathbb{E}_\varepsilon\big[h_n\big] := \mathbb{E}_\varepsilon\left[\frac{1}{\mu^n} \sum_{(i_1, \cdots, i_n)} h_n\big(\mathbf{Z}_{\underline{i}_n}\big)\right]$$

and the associated centered function defined on $\mathscr{D}_\varepsilon$

$$\hat{h}_n(\mathbf{Z}_\mathscr{N}) := \frac{1}{\mu^n} \sum_{(i_1, \cdots, i_n)} h_n\big(\mathbf{Z}_{\underline{i}_n}\big) - \mathbb{E}_\varepsilon\big[h_n\big].$$

2.6. **Strategy of the proof of Theorem 1.** The general strategy of proof of Theorem 1 is inspired by [BGSRS23b, LB25, LB23], in which the hard sphere system is studied. It is known, however, that the treatment of smooth potentials in the low-density limit leads to a number of delicate complications and methodology. This is already true for the law of large numbers, see eg [GSRT13, PSS14] where the proof of [Kin75] was completed and extended. The difficulties are due to the fact that interactions are not instantaneous. On the other hand, the extension to a long time of (2.15) for hard spheres, achieved in [LB25] (the short-time version was known long ago in [Spo83]) involves a different class of difficulty and technology. The main task of this paper is to push much further all the above-mentioned techniques in such a way as to allow the grazing collision limit $\alpha \to 0$. This combination is nontrivial, and we outline the main features and novelties involved to achieve the result.

As $\zeta_\varepsilon^0(g)$ is a mean-free random variable on $\mathscr{D}$, we can write

$$\mathbb{E}_\varepsilon\left[\zeta_\varepsilon^t(h)\zeta_\varepsilon^0(g)\right] = \frac{1}{\sqrt{\mu}}\mathbb{E}_\varepsilon\left[\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t))\zeta_\varepsilon^0(g)\right]. \tag{2.25}$$

We see that the function h is evaluated at time t, whereas the function g is evaluated at time 0. The first step of the proof is the construction of a family of functionals $(\Phi_{1,n})_{1\leq n}$, $\Phi_{1,n}^t : L^\infty(\mathbb{D}) \to L^\infty(\mathbb{D}^n)$ such that for any initial configuration $\mathbf{Z}_\mathscr{N} \in \mathscr{D}$,

$$\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t)) = \sum_{n\geq 1}\sum_{\underline{i}_n} \Phi_{1,n}^t[h](\mathbf{Z}_{\underline{i}_n}(0)). \tag{2.26}$$

The first part of Section 3 is dedicated to giving an explicit expression to the $\Phi_{1,n}^t$, *via* a *dynamical cluster development* (introduced by Sinai [Sin72] in a different setting and later by Bodineau *et al.* for

[6]from Edmund Landau and not Lev Landau.

hard spheres in Boltzmann-Grad scaling (see [BGSRS22])). If the particle 1 does not meet any other particles, it moves along a straight line. We deduce that

$$\Phi^t_{1,1}[h](x_1, v_1) = h(x_1 + tv_1, v_1).$$

If the particle 1 meets only the particle 2, then the trajectory of 1 (denoted $\mathsf{z}_1(t, 1 \text{ and } 2 \text{ interact})$) is governed by the Hamiltonian equation associated to the energy

$$\frac{|\mathsf{v}_1|^2}{2} + \frac{|\mathsf{v}_2|^2}{2} + \frac{\alpha}{2}\mathscr{V}\left(\frac{\mathsf{x}_1 - \mathsf{x}_2}{\varepsilon}\right),$$

with initial conditions $(\mathsf{z}_1(0), \mathsf{z}_2(0)) = (z_1, z_2)$. In order to write $\Phi^t_{1,2}[h](z_1, z_2)$, one also needs to compensate for the term $\Phi^t_{1,}[h](z_1)$ which appears now to be an error. Note that in this error term, the particles 1 and 2 do not interact. One writes

$$(\mathsf{z}_1, \mathsf{z}_2)(t, 1 \text{ and } 2 \text{ do not interact}) := (x_1 + tv_1, v_1, x_2 + tv_2, v_2).$$

We deduce that

$$\Phi^t_{1,2}[h](z_1, z_2) = [h(\mathsf{z}_1(t, 1 \text{ and } 2 \text{ interact}) - h(\mathsf{z}_1(t, 1 \text{ and } 2 \text{ do not interact}])\mathbb{1}_{1 \text{ and } 2 \text{ meet}}.$$

In order to generalize this construction, we define the pseudotrajectory $\mathsf{Z}_n(t, \text{history})$ (for the moment in a vague sense, see Definition 3.3 for a proper definition) as a trajectory of n particles where some couple of particles interacts and some couple of particles ignores each other. The "histories" are combinatorial parameters describing which couple of particle interacts (see Remark 3.4). We denote $\mathfrak{H}_n$ the set of histories (it will be precisely defined in Remark 3.4). We finally obtain

$$\Phi^t_{1,n}[h](Z_n) := \frac{1}{(n-1)!}\sum_{\text{history}\in\mathfrak{H}_n} h(\mathsf{z}_1(t, Z_n, \text{history}))\mathbb{1}_{\text{history}}\sigma(\text{history}) \tag{2.27}$$

In the preceding formula, $\mathbb{1}_{\text{history}}$ requires that the pseudotrajectory is possible (if we have all the needed collisions), and $\sigma(\text{history}) = \pm 1$ is a sign.

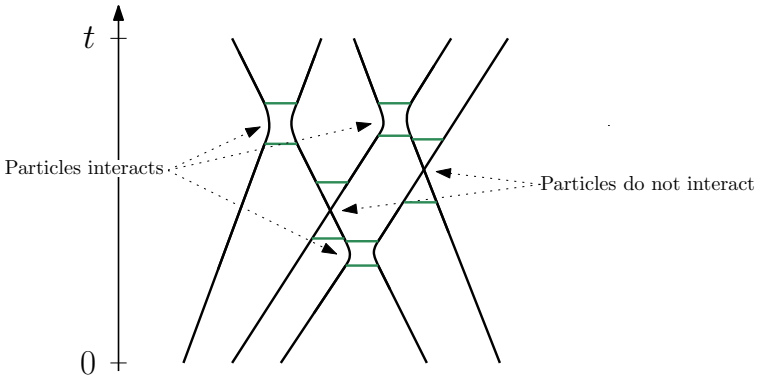

FIGURE 3. Exemple of pseudotrajectory for four particles.

Remark 2.6. *This representation differs from the one used in* [Kin75, GSRT13, PSS14]*, which is based on the Grad's representation of the BBGKY hierarchy* [Gra49]*. It leads to a drastic improvement to the argument in* [LB25] *(see below and Remark 3.8).*

Following the classical derivation of the Boltzmann equation (and here of the linearized Boltzmann equation), there will be two main steps. First, we need to prove that each term of the expansion in (2.26)

$$\mathbb{E}_\varepsilon\left[\mu^{-\frac{1}{2}}\sum_{\underline{i}_n}\Phi^t_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(0))\zeta^0_\varepsilon(g)\right] \tag{2.28}$$

converges to its formal limit. In the limit, the dynamics loses its Hamiltonian character, and particles become punctual (see Section 6).

The main obstacles to this convergence are *multiple encounters* (interactions between more than three particles) and *recollisions*. A recollision can be defined (a more proper definition will be given in Definition 3.7) as a meeting (with interaction or with overlap) between two particles q and $\bar{q}$, beginning at time τ and such that we can find a sequence of couples of particles $(q = q_1, \bar{q}_1), (q_2, \bar{q}_2) \cdots (q_r, \bar{q}_r = \bar{q})$,

such that $q_{i+1} \in \{q_i, \bar{q}_i\}$ and that q_i and $\bar{q}_i$ meet before time τ. Recollisions and multiple encounters become rare in the limit $\varepsilon \to 0$ (quantitative estimations are performed in Section B.2).

The second step is an *a priori* bound of the terms of the series (2.26). To improve estimates to a longer time interval than the one obtained in [GSRT13, PSS14], it is convenient to consider L^2 estimates (see [BGSR17, BGSRS23b, BGSRS24, LB25]). Indeed, because $\zeta_\varepsilon^0(g)$ is a mean-free random variable, for any intermediate time $t_s \in [0,t]$,

$$\begin{aligned}\left|\mathbb{E}_\varepsilon\left[\mu^{-\frac{1}{2}}\sum_{\underline{i}_n}\Phi_{1,n}^{t-t_s}[h](\mathbf{Z}_{\underline{i}_n}(t_s))\zeta_\varepsilon^0(g)\right]\right| &= \left|\mathbb{E}_\varepsilon\left[\mu^{n-\frac{1}{2}}\hat{\Phi}_{1,n}^{t-t_s}[h](\mathbf{Z}_{\mathscr{N}}(t_s))\zeta_\varepsilon^0(g)\right]\right| \\ &\le \mathbb{E}_\varepsilon\left[\mu^{2n-1}\left(\hat{\Phi}_{1,n}^{t-t_s}[h](\mathbf{Z}_{\mathscr{N}}(t_s))\right)^2\right]^{\frac{1}{2}}\mathbb{E}_\varepsilon\left[\left(\zeta_\varepsilon^0(g)\right)^2\right]^{\frac{1}{2}} \\ &\le \mathbb{E}_\varepsilon\left[\mu^{2n-1}\left(\hat{\Phi}_{1,n}^{t-t_s}[h](\mathbf{Z}_{\mathscr{N}}(0))\right)^2\right]^{\frac{1}{2}}\mathbb{E}_\varepsilon\left[\left(\zeta_\varepsilon^0(g)\right)^2\right]^{\frac{1}{2}},\end{aligned}$$

using a Cauchy-Schwartz inequality and the invariance of the Gibbs measure. Hence, it is possible to start a development along pseudotrajectories and stop at time t_s when they become "pathological". We say that pseudotrajectories become pathological when their number explodes (this notion will be made precise after (2.33) and is linked to recollisions and multiple encounters). Then we can ignore what happens in the time interval $[0,t_s]$.

We need to bound the $\mathbb{E}_\varepsilon\left[\left(\hat{\Phi}_{1,n}^t[h]\right)^2\right]$, which are based on the estimations of terms of type

$$\sup_{l\le n}\frac{n^l}{\mu^{l-1}}\int\left|\Phi_{1,n}^{t-t_s}[h](Z_n)\Phi_{1,n}^{t-t_s}[h](Z_{[n-l,2n-l]})\right|e^{-\mathscr{H}_{2n-l}(Z_{2n-m})}\,\mathrm{d}Z_{2n-m}, \tag{2.29}$$

see Proposition 4.1.

Remark 2.7. *An $L^\infty \to L^1$ estimation is used in the classical derivation of the Boltzmann equation (see* [Lan75, Kin75, GSRT13, PSS14]*). It is valid only for short times. The linear version of the problem (one tagged particle followed in a background initially at equilibrium) is only a $O(1)$ perturbation of equilibrium in L^∞. Thus, the L^1 bounds are valid for all time (see* [vBLLS80, BGSR16, Ayi17, Cat18]*). The linearized setting is a $O(\mu)$ perturbation of the equilibrium, and L^1 bounds are no longer sufficient to reach long time out of equilibrium (Spohn used them to describe the fluctuations on short time in* [Spo81]*).*

Unfortunately, we do not know how to efficiently take into account the sign σ(history) in the bound of $\Phi_{1,n}^{t-t_s}$. Thus, we resort to using the naive bound

$$\left|\Phi_{1,n}^t[h]\right| \lesssim \frac{\|h\|}{(n-1)!}\sum_{\text{history}\in\mathfrak{H}_n'}\mathbb{1}_{\text{history}}. \tag{2.30}$$

where $\mathfrak{H}_n' \subset \mathfrak{H}_n$ is a set of parameters that is a little bit smaller than $\mathfrak{H}_n$ (it will be made precise in Remark 3.4). Two particles with initial coordinates (x_1,v_1) and (x_2,v_2) can collide on the time interval $[0,t]$ if and only if x_2 is in a cylinder of radius ε and length $t|v_1-v_2|$. We deduce that

$$\int\mathbb{1}_{1\text{ and }2\text{ collide on }[0,t]}e^{-\mathscr{H}_2}\,\mathrm{d}Z_2 \lesssim t\varepsilon^{d-1} \lesssim \frac{t}{\mu\mathfrak{d}}. \tag{2.31}$$

For a pseudotrajectory involving n particles, there are at least $n-1$ collisions (all particles are linked by a chain of collisions). Using a generalization of the preceding argument, we deduce the following L^1 bound

$$\int\sum_{\text{history}\in\mathfrak{H}_n'}\mathbb{1}_{\text{history}\in\mathfrak{H}_n'}e^{-\mathscr{H}_n}\,\mathrm{d}Z_n \lesssim n^n\left(\frac{t}{\mu\mathfrak{d}}\right)^{n-1} \tag{2.32}$$

for some C (a rigorous proof is provided in Proposition II.6 of [BGSRS22][7]). As we perform L^2 estimates, we will encounter terms of the form $\mathbb{1}_{\text{history}_1}(Z_n)\mathbb{1}_{\text{history}_2}(Z_n)$ where $(\text{history}_1,\text{history}_2)\in(\mathfrak{H}_n')^2$ are two different histories. The problem is that the geometry of the set $\{Z_n|\mathbb{1}_{\text{history}_1}(Z_n)=1\}$ is quite complicated and we do not know a better estimate than the naive bound

$$\mathbb{1}_{\text{history}_1}(Z_n)\mathbb{1}_{\text{history}_2}(Z_n) \le \mathbb{1}_{\text{history}_1}(Z_n).$$

[7]The estimation is provided in [BGSRS22] in the hard sphere setting but should be generalized to more general interaction. We will not need this precise estimation, and we will limit ourselves to proving for a given history

$$\int\mathbb{1}_{\text{history}}e^{-\mathscr{H}_n}\,\mathrm{d}Z_n \lesssim \left(\tfrac{Cnt}{\mu\mathfrak{d}}\right)^{n-1}.$$

From this, we may obtain the bound on (2.29) bigger than

$$\frac{n^n}{\mu^{n-1}} \int \left|\Phi^t_{1,n}[h]\right|^2 e^{-\mathscr{H}_n} \, \mathrm{d}Z_n \lesssim \|h\| \left(\frac{Ct}{\mu^2 \mathfrak{d}}\right)^{n-1} |\mathfrak{H}'_n| \,. \tag{2.33}$$

Thus, we are led to counting the number of "histories", *id est*, the number of pseudotrajectories involving n particles. The set $|\mathfrak{H}'_n|$ is too large (it is of order $O((Cn)^n)$ for some constant $C > 0$, see Remark 3.4) to allow the series $\sum_n \mathbb{E}\left[(\hat{\Phi}^t_{1,n})^2\right]^{\frac{1}{2}}$ to converge. We would need that $|\mathfrak{H}'_n|$ is of order $O(C^n)$ for some constant $C > 0$. To overcome this problem, we will perform the pseudotrajectory decomposition only until an intermediate time t_s, such that the set of collision parameters needed $\mathfrak{H}''_n$ remains controllable. We observe that there are two reasons for the set of parameters to explode: the *pathology* (recollisions or multiple meetings) and an uncontrolled number of particles n.

We now introduce two samplings, one to control regular collisions and one to control recollisions and multiple encounters .

The first sampling has a relatively large step $\theta := \frac{1}{\beta \log|\log \varepsilon|}$ (for some constant β large enough). We stop the pseudotrajectory development at time $t - k\theta$ if there are more than 2^k particles involved in the pseudotrajectory. Hence, the number of particles at time 0 remains controlled. This sampling follows [BGSR16, BGSR17][8].

The second sampling has a shorter step, $\delta := \varepsilon^{1/12}$. We stop the expansion at time $t_s := t - k\delta$ if the pseudotrajectory has at least one recollision on $[t_s, t]$ (but no recollision on $[t_s + \delta, t]$). Imposing recollisions creates an additional geometric condition, and thus, an extra-smallness gain.

Unfortunately, we still have too many possible histories. In order to reduce their number, we follow the idea of [LB25] and separate the pseudotrajectories into two categories. In *non-pathological* pseudotrajectories, there is no recollision nor multiple encounters on the time interval $[t_s, t_s + \delta]$. We are in a setting close to the case without recollision, and we only need C^n parameters (C a fixed constant) to describe the histories.

We explain now how to treat the pathological recollisions part. The initial data $\mathbf{Z}_{\mathscr{N}}(0)$ is conditioned such that on each interval $[k\delta, (k+1)\delta]$ a particle can *encounter*[9] with only a finite number of particles γ (we will take $\gamma := 12d$). Hence, for a pseudotrajectory $\mathrm{z}_1(t, Z_n, \text{history})$, the history has to describe first a partition of $[n]$ into small clusters of particles that interact together on $[0, \delta]$ and how they really interact. As the size of each cluster is uniformly bounded, the number of histories is at most of order C^n for some $C > 1$.

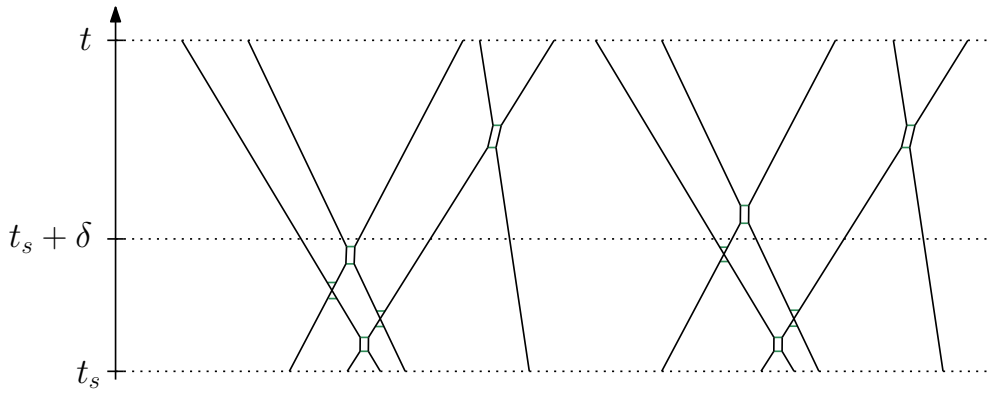

FIGURE 4. An example of one pathological pseudotrajectory (on the left) and a non-pathological one (on the right)

The paper is organised as follows:

- In Section 3, we give a proper definition of histories, and we use it to construct the functionals $\Phi^t_{1,n}$. Then we implement the two samplings. This allows us to decompose $\mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(h)\zeta^0_\varepsilon(g)\right]$ into a main term, plus error terms.

[8]Note that in [Fou24], the author improved the sampling strategy in the linear setting in order to obtain a better convergence rate than (2.15).

[9]The meaning of *encounter* will be precise in Definition 3.5.

- In Section 4, we introduce standard $L^2(\mathbb{P}_\varepsilon)$ estimates that are necessary to bound terms of the form $\mathbb{E}_\varepsilon\left|\hat{\Psi}\right|^2$ (for some symmetric function $\Psi_n : \mathbb{D}^n \to \mathbb{R}$) by the integrals
$$\forall m \leq n,\ \int \left|\Psi_n(Z_n)\Psi_n(Z_{[n-m,2n-m]})\right| e^{-\mathscr{H}_{2n-m}(Z_{2n-m})}\,\mathrm{d}Z_{2n-m}.$$
The proof is an adaptation of Section 3 of [BGSRS23b], and is based on static cumulant decompositions.
- Section 5 is dedicated to the bound of the pseudotrajectory development $\Phi^0_{1,n}$ in the case where recollisions and multiple meetings are forbidden. This allows us to treat the pseudotrajectory with many particles.
- Section 6 is dedicated to the convergence of the main term.
- Section 7 is dedicated to the treatment of the non-pathological recollisions.
- Section 8 is dedicated to the treatment of pathological recollisions and of multiple meetings.
- Annex B is dedicated to the analyses of trajectories leading to recollisions of multiple encounters . We use a strategy similar to the one of [PSS14]. It has the advantage of giving estimations independent of the form of the potential $\mathscr{V}$ (supposing it respects Assumption 2.1).

3. Development along pseudotrajectories and time sampling

3.1. **Dynamical cluster development.** For any test functions h and $g : \mathbb{D} \to \mathbb{R}$, we want to compute
$$\mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(h)\zeta^0_\varepsilon(g)\right] = \frac{1}{\mu}\mathbb{E}_\varepsilon\left[\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t))\sum_{j=1}^{\mathscr{N}} g(\mathbf{z}_j(0))\right].$$

We have a sum evaluated at time t and a sum evaluated at time 0. In order to compute it, we have to pull back the second sum to time 0: we want to construct a family of applications $\Phi^t_{1,n} : L^\infty(\mathbb{D}) \to L^\infty(\mathbb{D}^n)$ such that for almost all initial data $\mathbf{Z}_{\mathscr{N}}(0) \in \mathscr{D}$
$$h(\mathbf{z}_{i_1}(t)) = \sum_{n\geq 1}\sum_{(i_2,\cdots,i_n)} \Phi^t_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(0)).$$
More generally, we will construct a family of functional $\Phi^t_{m,n} : L^\infty(\mathbb{D}^m) \to L^\infty(\mathbb{D}^n)$ (with $m < n$) such that for any test functions $h_m \in L^\infty(\mathbb{D}^m)$,
$$h_m(\mathbf{Z}_{\underline{i}_m}(t)) = \sum_{n\geq 1}\sum_{(i_{m+1},\cdots,i_n)} \Phi^t_{m,n}[h_m](\mathbf{Z}_{\underline{i}_n}(0)).$$

Remark 3.1 (Comparison with the hard sphere setting)**.** *In the hard spheres setting, a tree pseudotrajectories development is used as it comes directly from the BBGKY hierarchy (see, for example,* [Lan75, PS15, BGSRS23b, LB25]*). We begin at time* 0 *with* n *particles, and at each collision, we can remove or not remove one particle to end at time* t *with* m *particles. However, in the case of physical potential, writing the BBGKY hierarchy is difficult as particles can overlap, and there can be interaction between more than three particles (see* [Gra49, Kin75, GSRT13, PSS14] *for a description of the BBGKY hierarchy). Hence, we will use a different kind of pseudotrajectory development called "dynamical cluster development" (see* [Sin72, PSW16, BGSRS22]*, from which we take inspiration).*

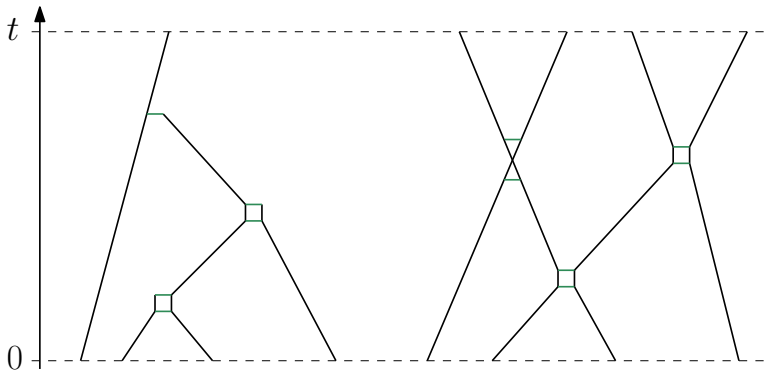

Figure 5. On the left a tree pseudotrajectory, on the right a graph pseudotrajectory.

Fix $\lambda \subset \mathbb{N}$ a finite set of particles. We denote by $\mathsf{Z}^\lambda(\tau) = (\mathsf{X}^\lambda(\tau), \mathsf{V}^\lambda(\tau))$ the Hamiltonian trajectory, linked to the energy

$$\mathscr{H}_\lambda(Z_\lambda) := \sum_{q\in\lambda} \frac{|v_q|^2}{2} + \frac{\alpha}{2} \sum_{\substack{q,\bar q\in\lambda \\ q\neq\bar q}} \mathscr{V}\left(\frac{x_q - x_{\bar q}}{\varepsilon}\right)$$

of the particles λ (isolated of the other particles) with initial data $\mathsf{Z}^\lambda(0) = Z_\lambda$. For any subset $\lambda' \subset \lambda$, we denote $\mathsf{Z}^\lambda_{\lambda'}(\tau)$ the trajectory of particles λ' in $\mathsf{Z}^\lambda(\tau)$.

Definition 3.1. *Given $Z_\lambda \in \mathbb{D}^{|\lambda|}$, we construct the graph G with vertex λ and $(q,\bar q) \in \lambda^2$ is an edge if and only if $q < \bar q$ and if there exists a time $\tau \in [0,t]$ such that*

$$\exists \tau \in [0,t],\ \left|\mathsf{x}^\lambda_q(\tau) - \mathsf{x}^\lambda_{\bar q}(\tau)\right| \leq \varepsilon.$$

We say that $\mathsf{Z}^\lambda(\tau)$ forms a dynamical cluster *if the graph G is connected. We denote $\mathbb{1}_{|\lambda|}(Z_\lambda)$ the indicator function that the trajectory $\mathsf{Z}^\lambda(\tau)$ forms a dynamical cluster.*

In the same way, for $\omega \subset \lambda$, we say that $\mathsf{Z}^\lambda(\tau)$ forms a ω-cluster *if, in the collision of $\mathsf{Z}^\lambda(\tau)$, all the particles are in the same connected components of G that one of the particles of ω. The function $\mathbb{1}^\omega_{|\lambda|}(Z_\lambda)$ is equal to 1 if $\mathsf{Z}^\lambda(\tau)$ is a ω-cluster, 0 else.*

Remark 3.2. *In the following, we consider that all the graphs are unoriented.*

Definition 3.2. *We say that trajectories $\mathsf{Z}^\lambda(\tau)$ and $\mathsf{Z}^{\lambda'}(\tau)$ (with $\lambda \cap \lambda' = \emptyset$) have a* dynamical overlap *if there exists a couple of particles $(q,q') \in \lambda \times \lambda'$ and some time $\tau \in [0,t]$, such that $|\mathsf{x}^\lambda_q(\tau) - \mathsf{x}^{\lambda'}_{q'}(\tau)| \leq \varepsilon$. Then we denote $\lambda \overset{\circ}{\sim} \lambda'$.*

For $(Z_{\lambda_1}, \cdots, Z_{\lambda_{\mathbf{l}}}) \in \prod_{i=1}^{\mathbf{l}} \mathbb{D}^{|\lambda_i|}$ initial data, we look at the indicator function that for any $i \neq j$, $\mathsf{Z}^{\lambda_i}(\tau)$ and $\mathsf{Z}^{\lambda_j}(\tau)$ have no dynamical overlap. We can expand it as

$$\prod_{1\leq i<j\leq \mathbf{l}} \left(1 - \mathbb{1}_{\lambda_i \overset{\circ}{\sim} \lambda_j}\right) = \sum_{\substack{\omega\subset[1,l] \\ 1\in\omega}} \underbrace{\sum_{C\in\mathscr{C}(\omega)} \prod_{(i,j)\in E(C)} -\mathbb{1}_{\lambda_i \overset{\circ}{\sim} \lambda_j}}_{:= \mathscr{O}_{|\omega|}(Z_{\lambda_1}, Z_{\lambda_{\omega(2)}}, \cdots, Z_{\lambda_{\omega(|\omega|)}})} \prod_{\substack{(i,j)\in([\mathbf{l}]\setminus\omega)^2 \\ i\neq j}} \left(1 - \mathbb{1}_{\lambda_i \overset{\circ}{\sim} \lambda_j}\right). \tag{3.1}$$

We have defined $(\mathscr{O}_l)_l$ as the cumulants of the dynamical overlap indicator functions.

We make a partition of $\mathscr{D}$ depending on the way particles interact during the time interval $[0,t]$: fixing $\mathscr{N} \in \mathbb{N}$ and $\underline{i}_m$, (we recall that $\mathscr{P}^r_\omega$ is the set of partitions into r subsets of ω)

$$\begin{aligned} h_m(\mathbf{Z}_{\underline{i}_m}(t)) &= \sum_{\mathbf{l}=1}^{\mathscr{N}} \sum_{\substack{\underline{i}_m\subset\lambda_1 \\ (\lambda_2,\cdots,\lambda_{\mathbf{l}})\in\mathscr{P}^{\mathbf{l}-1}_{[\mathscr{N}]\setminus\lambda_1}}} h_m(\mathbf{Z}_{\underline{i}_m}(t)) \mathbb{1}^{\underline{i}_m}_{\lambda_1}(\mathbf{Z}_{\lambda_1}) \prod_{i=2}^{\mathbf{l}} \mathbb{1}_{|\lambda_i|}(\mathbf{Z}_{\lambda_i}) \prod_{1\leq i<j\leq\mathbf{l}} \left(1 - \mathbb{1}_{\lambda_i \overset{\circ}{\sim} \lambda_j}\right) \\ &= \sum_{\mathbf{l}=1}^{\mathscr{N}} \sum_{\substack{\underline{i}_m\subset\lambda_1 \\ (\lambda_2,\cdots,\lambda_{\mathbf{l}})\in\mathscr{P}^{\mathbf{l}-1}_{[\mathscr{N}]\setminus\lambda_1}}} h_m(\mathbf{Z}_{\underline{i}_m}(t)) \mathbb{1}^{\underline{i}_m}_{\lambda_1}(\mathbf{Z}_{\lambda_1}) \prod_{i=2}^{\mathbf{l}} \mathbb{1}_{|\lambda_i|}(\mathbf{Z}_{\lambda_i}) \sum_{\substack{\omega\subset[\mathbf{l}] \\ 1\in\omega}} \mathscr{O}_{|\omega|}(\mathbf{Z}_{\underline{\lambda}_\omega}) \\ &\qquad \times \prod_{\substack{(i,j)\in(\omega\setminus[\mathbf{l}])^2 \\ i\neq j}} \left(1 - \mathbb{1}_{\lambda_i \overset{\circ}{\sim} \lambda_j}\right). \end{aligned}$$

We make the change of variables

$$(\mathbf{l}, (\lambda_1, \cdots, \lambda_{\mathbf{l}}), \omega) \mapsto \left(\rho, \mathbf{l}_1, \left(\bar\lambda_1, \cdots, \bar\lambda_{\mathbf{l}_1}\right), \mathbf{l}_2, \left(\tilde\lambda_1, \cdots, \tilde\lambda_{\mathbf{l}_2}\right)\right)$$

where

$$\rho := \bigcup_{i\in\omega} \lambda_i,\ \mathbf{l}_1 := |\omega|,\ \mathbf{l}_2 := \mathbf{l} - |\omega|,\ \bar\lambda_1 := \lambda_1,$$

$$\left(\bar\lambda_2, \cdots, \bar\lambda_{\mathbf{l}_1}\right) := (\lambda_j)_{j\in\omega\setminus\{1\}} \text{ and } \left(\tilde\lambda_1, \cdots, \tilde\lambda_{\mathbf{l}_2}\right) := (\lambda_j)_{j\in[\mathbf{l}]\setminus\omega}.$$

The set ρ is the set of particles linked to $\underline{i}_m$ *via* a chain of interactions or overlaps. We get that $h_m(\mathbf{Z}_{\underline{i}_m}(t))$ is equal to

$$\sum_{\rho \supset \underline{i}_m} \sum_{\mathbf{l}_1=1}^{|\rho|} \sum_{\substack{\underline{i}_m \subset \bar{\lambda}_1 \subset \rho \\ (\bar{\lambda}_2, \cdots, \bar{\lambda}_{\mathbf{l}_1}) \in \mathscr{P}_{\rho \setminus \bar{\lambda}_1}^{\mathbf{l}_1 - 1}}} h_m\left(\mathsf{Z}_{\underline{i}_m}^{\bar{\lambda}_1}(t)\right) \mathbb{M}_{\bar{\lambda}_1}^{\underline{i}_m}\left(\mathbf{Z}_{\bar{\lambda}_1}\right) \prod_{i=2}^{\mathbf{l}_1} \mathbb{M}_{|\bar{\lambda}_i|}\left(\mathbf{Z}_{\bar{\lambda}_i}\right) \mathscr{O}_{\mathbf{l}_1}\left(\mathbf{Z}_{\bar{\lambda}_1}, \cdots, \mathbf{Z}_{\bar{\lambda}_{\mathbf{l}_1}}\right)$$
$$\times \sum_{\mathbf{l}_2=1}^{\mathscr{N}-|\rho|} \sum_{(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_{\mathbf{l}_2}) \in \mathscr{P}_{[\mathscr{N}] \setminus \rho}^{\mathbf{l}_2}} \prod_{i=1}^{\mathbf{l}_2} \mathbb{M}_{|\tilde{\lambda}_i|}(\mathbf{Z}_{\tilde{\lambda}_i}) \prod_{\substack{(i,j) \in ([\mathscr{N}] \setminus \omega)^2 \\ i \neq j}} \left(1 - \mathbb{1}_{\tilde{\lambda}_i \overset{\circ}{\sim} \tilde{\lambda}_j}\right).$$

The second line is the sum over all possible partitions $(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_{\mathbf{l}_2})$ of $[\mathscr{N}] \setminus \rho$ of the indicator function that they are effectively the dynamical cluster of the initial data. Hence, it is equal to one. We identify the n-th *dynamical cumulant* as

$$\Phi_{m,n}^t[h_m](Z_n) := \frac{1}{(n-m)!} \sum_{\mathbf{l}=1}^{n} \sum_{[m] \subset \lambda_1 \subset [n]} \sum_{\substack{(\lambda_2, \cdots, \lambda_{\mathbf{l}}) \\ \in \mathscr{P}_{[n] \setminus \lambda_1}^{\mathbf{l}-1}}} h_m(\mathsf{Z}_{[m]}^{\lambda_1}(t)) \mathbb{M}_{|\lambda_1|}^{[m]}(Z_{\lambda_1}) \prod_{i=2}^{\mathbf{l}} \mathbb{M}_{|\lambda_i|}(Z_{\lambda_i}) \times \mathscr{O}_l(Z_{\lambda_1}, \cdots, Z_{\lambda_{\mathbf{l}}}), \tag{3.2}$$

and we can now write the dynamical cluster expansion:

Proposition 3.1. *Fix a family of particle $\underline{i}_m$. For almost all $\mathbf{Z}_{\mathscr{N}} \in \mathscr{D}$, we have*

$$h_m\left(\mathbf{Z}_{\underline{i}_m}(t)\right) = \sum_{n \geq m} \sum_{(i_{m+1}, \cdots, i_n)} \Phi_{m,n}^t[h_m]\left(\mathbf{Z}_{\underline{i}_n}(0)\right). \tag{3.3}$$

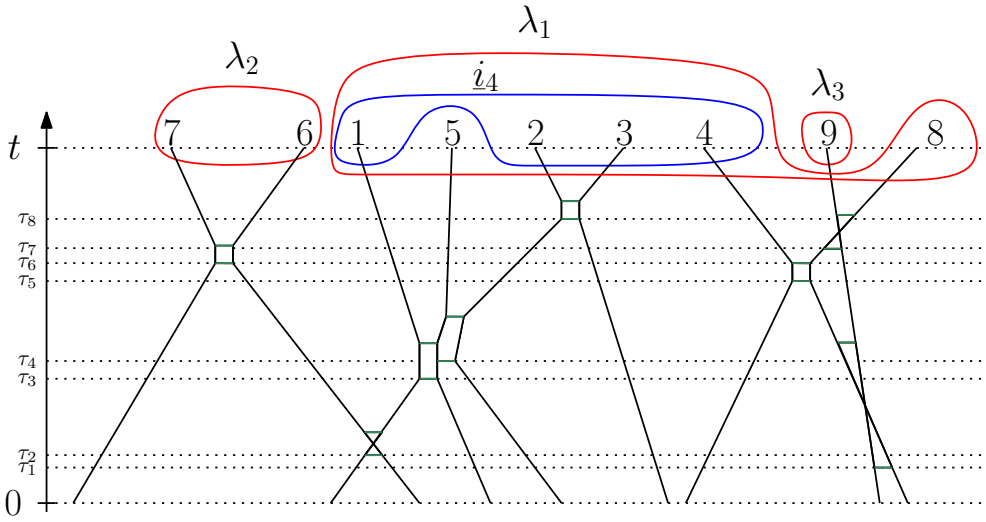

FIGURE 6. Example of trajectory in a dynamical cumulant. We want to follow the particles $\{1, 2, 3, 4\}$.

Definition 3.3 (First type of pseudotrajectory)**.** *In the following, for a given $m \leq n$, $\underline{\lambda} = (\lambda_1, \cdots, \lambda_{\mathbf{l}})$ a partition of $[n]$, we denote $\mathsf{Z}(t, Z_n, \underline{\lambda})$ the trajectory of the n particles following the Hamiltonian dynamics linked to*

$$\mathscr{H}_{\underline{\lambda}}(Z_n) := \sum_{\ell=1}^{\mathbf{l}} \mathscr{H}_{\lambda_\ell}(Z_{\lambda_\ell}).$$

We define now the notion of collision graph:

Definition 3.4 (collision graph)**.** *Fix $m \leq n$, collision parameters $\underline{\lambda} := (\lambda_1, \cdots, \lambda_{\mathbf{l}})$ and an initial position $Z_n \in \mathbb{D}^n$.*

We construct the collision graph *with vertex $[n]$ and with labeled edges of the form $(i,j)_{\tau,s}$, $\tau \in [0,t]$, $s \in \{\pm 1\}$. The edges $(i,j)_{\tau,s}$ is in the graph if*

- $\tau \in (0,t)$, $|\mathsf{x}_i(\tau, \underline{\lambda}) - \mathsf{x}_j(\tau, \underline{\lambda})| = \varepsilon$, $(\mathsf{x}_i(\tau, \underline{\lambda}) - \mathsf{x}_j(\tau, \underline{\lambda})) \cdot (\mathsf{v}_i(\tau, \underline{\lambda}) - \mathsf{v}_j(\tau, \underline{\lambda})) > 0$,

- *or* $\tau = 0$, $\left|\mathsf{x}_i(0,\bar\lambda) - \mathsf{x}_j(0,\bar\lambda)\right| < \varepsilon$,
- $s = 1$ *if* i *and* j *are in the same* λ_k, $s = -1$ *else.*

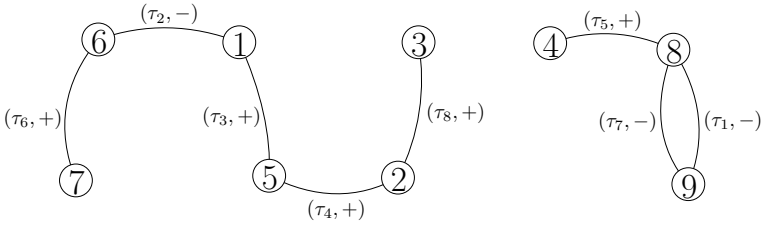

FIGURE 7. The collision graph associated with the pseudotrajectory of Figure 3.1.

Remark 3.3. *Fix* $(\lambda_1, \cdots, \lambda_\ell)$ *a partition of* $[n]$. *Using Penrose's tree inequality (see* [Pen63, BGSRS20, Jan]*), the cumulant function* $\mathscr{O}_n(Z_{\lambda_1}, \cdots, Z_{\lambda_n})$ *is bounded by*

$$|\mathscr{O}_n(Z_{\lambda_1}, \cdots, Z_{\lambda_n})| \le \sum_{T \in \mathscr{T}([\ell])} \prod_{(i,j)\in E(T)} \mathbb{1}_{\lambda_i \mathring{\sim} \lambda_j} \tag{3.4}$$

where $\mathscr{T}([\ell])$ *is the set of simply connected graph on* $[\ell]$. *The case of equality is reached, so we cannot expect a good* L^∞ *bound of* $\mathscr{O}_n$.

Remark 3.4. *We will describe the set* $\{history\}$ *that has been used in Section 2.6 for the decomposition of* $\Phi^t_{1,n}[h]$*:* history *is a triplet* $(\mathbf{l}, (\lambda_2, \cdots, \lambda_{\mathbf{l}}), C) \in \mathfrak{H}_n$, *where the set of history* $\mathfrak{H}_n$ *is defined by*

$$\mathfrak{H}_n := \left\{ (\mathbf{l}, (\lambda_2, \cdots, \lambda_{\mathbf{l}}), C) \middle| \mathbf{l} \in [1,n]\,;\, \{1\} \subset \lambda_1 \subset [n]\,;\, (\lambda_2, \cdots, \lambda_{\mathbf{l}}) \in \mathscr{P}^{\mathbf{l}-1}_{\lambda_1^c}\,;\, C \in \mathscr{C}([\mathbf{l}]) \right\},$$

$$\mathbb{1}_{\text{history}} := \prod_{i=1}^{\mathbf{l}} \mathbb{M}_{|\lambda_i|}(\mathbf{Z}_{\lambda_i}) \prod_{(i,j)\in E(C)} \mathbb{1}_{\lambda_i \mathring{\sim} \lambda_j}$$

$$\sigma(\text{history}) := (-1)^{|E(C)|},$$

where $|E(C)|$ *is the number of edges of the graph* C. *One has*

$$\Phi^t_{1,n}[h](Z_n) := \frac{1}{(n-1)!} \sum_{\text{history}\in\mathfrak{H}_n} h(\mathsf{z}_1^{\lambda_1}(t, Z_{\lambda_1}) \mathbb{1}_{\text{history}} \sigma(\text{history}).$$

Using the Penrose's tree inequality (3.4)*, one obtains*

$$\left|\Phi^t_{1,n}[h](Z_n)\right| \le \frac{1}{(n-1)!} \sum_{\text{history}\in\mathfrak{H}'_n} h(\mathsf{z}_1^{\lambda_1}(t, Z_{\lambda_1}) \mathbb{1}_{\text{history}} \sigma(\text{history})$$

where $\mathfrak{H}'_n \subset \mathfrak{H}_n$ *is a smaller set of parameters*

$$\mathfrak{H}'_n := \left\{ (\mathbf{l}, (\lambda_2, \cdots, \lambda_{\mathbf{l}}), C) \middle| \mathbf{l} \in [1,n]\,;\, \{1\} \subset \lambda_1 \subset [n]\,;\, (\lambda_2, \cdots, \lambda_{\mathbf{l}}) \in \mathscr{P}^{\mathbf{l}-1}_{[n]\setminus\lambda_1}\,;\, C \in \mathscr{T}([\mathbf{l}]) \right\}.$$

Using that $|\cup_{\mathbf{l}} \mathscr{P}^{\mathbf{l}}_{[n]}| \le e^n$, $|\mathscr{T}([\mathbf{l}])| = \mathbf{l}^{\mathbf{l}-2}$,

$$\begin{aligned} |\mathfrak{H}'_n| &= \sum_{\mathbf{l}=1}^{n} \left| \left\{ (\mathbf{l}, (\lambda_2, \cdots, \lambda_{\mathbf{l}})) \middle| \mathbf{l} \in [1,n]\,;\, \{1\} \subset \lambda_1 \subset [n]\,;\, (\lambda_2, \cdots, \lambda_{\mathbf{l}}) \in \mathscr{P}^{\mathbf{l}-1}_{[n]\setminus\lambda_1} \right\} \right| \times \left|\mathscr{T}([\mathbf{l}])\right|, \\ &\le \sum_{\mathbf{l}=1}^{n} C^n \times \mathbf{l}^{\mathbf{l}-2} \le C^n n^n \end{aligned}$$

for some constant $C > 0$, *which is faster in* n *than any geometric sequence. This shows that estimations of the form* (2.33) *cannot be useful yet.*

3.2. **Conditioning.** We describe now the conditioning used to control the pathological recollisions, that were introduced in Section 2.6.

Definition 3.5 (Possible cluster)**.** *Let $Z_r \in \mathbb{D}^r$ an initial configuration. Consider $\omega_1, \cdots, \omega_p$ a family of subsets of $[r]$ such that*

$$\bigcup_{i=1}^{p} \omega_i = [r],$$

and $(\underline{\lambda}_i)_{i\leq p} = (\lambda_i^1, \cdots, \lambda_i^{l_i})_{i\leq p}$ where each $\underline{\lambda}_i$ is a partition of the corresponding ω_i. We denote $\mathscr{G}_i$ the collision graph of the pseudotrajectory $\mathsf{Z}(\tau, Z_{\omega_i}, \underline{\lambda}_i)$ on the time interval $[0,\delta]$. The graph $\mathscr{G}$ is the merge of all the $\mathscr{G}_i$.

We say that Z_r forms a possible cluster *if there exists a couple $((\omega_i)_i, (\underline{\lambda}_i)_i)$ such that the graph $\mathscr{G}$ is connected.*

Definition 3.6 (Definition of the set Υ_ε)**.** *Let $\gamma > 0$ be a fixed integer, $\delta > 0$ a time scale, and $\mathbb{V} > 0$ a velocity bound. We construct $\Upsilon_\varepsilon \subset \mathscr{D}$ the set of particle configurations such that for any time $\tau \in \{0, \delta, 2\delta, \cdots, t\}$, there is no possible cluster of size bigger than γ at time τ, and inside any subset of particles $\omega \subset [1, \mathscr{N}]$ with less than γ elements, $\frac{1}{2}\|\mathbf{V}_\omega(\tau)\|^2$ is bounded by $\frac{1}{2}\mathbb{V}^2$.*

We have the following bound on the measure of the complement of Υ_ε:

Proposition 3.2. *There exists a constant C_γ depending only on γ and on the dimension such that*

$$\mathbb{P}_\varepsilon\left(\Upsilon_\varepsilon^c\right) \leq C_\gamma \frac{t}{\delta}\left(\mu\delta^\gamma + \mu^\gamma e^{-\mathbb{V}^2/4}\right). \tag{3.5}$$

Proof. We take the notation of the definition 3.5.

Suppose that Z_r (with $r > \gamma$) forms a possible cluster. We want to show that there exists a subset $\varpi \subset [r]$ such that Z_ϖ forms a possible cluster and $\gamma + 1 \leq |\varpi| \leq 2\gamma + 2$. As Z_r forms a possible cluster, there exists a couple $((\omega_i)_i, (\underline{\lambda}_i)_i)$ such that the graph $\mathscr{G}$ defined in Definition 3.6 is connected. We define $\mathscr{G}_\tau$ as the subgraph of $\mathscr{G}$ with edges

$$\left\{(q, \bar{q})_{(\tau',\sigma)} \in E(\mathscr{G}), \tau' < \tau\right\}.$$

We consider τ the infimum of $\{\tau' > 0, \mathscr{G}_{\tau'} \text{ is connected}\}$. Then, the graph $\mathscr{G}_\tau$ has exactly two connected components. One of them, denoted ϖ, verifies $\lceil \frac{r}{2} \rceil \leq |\varpi| \leq r - 1$. Iterating the procedure, we obtain the expected result.

We deduce that

$$\begin{aligned}\mathbb{P}_\varepsilon\left(\Upsilon_\varepsilon^c\right) &\leq \sum_{k=0}^{t/\delta} \mathbb{E}_\varepsilon \left[\sum_{n=\gamma+1}^{2(\gamma+1)} \frac{1}{n!} \sum_{(i_1,\cdots,i_n)} \mathbb{1}_{\substack{\mathbf{Z}_{i_n}(k\delta) \text{ form a} \\ \text{possible cluster}}} + \sum_{n=1}^{\gamma} \frac{1}{n!} \sum_{(i_1,\cdots,i_n)} \mathbb{1}_{\|\mathbf{V}_{i_n}(k\delta)\| \geq \mathbb{V}}\right] \\ &\leq \frac{t}{\delta}\left(\sum_{n=\gamma+1}^{2(\gamma+1)} \frac{1}{n!}\mu^n \int \mathbb{1}_{\substack{Z_n \text{ form a} \\ \text{possible cluster}}} M^{\otimes n}\, \mathrm{d}Z_n + \sum_{n=1}^{\gamma} \frac{1}{n!}\mu^n \int \mathbb{1}_{\|V_{k'}\| \geq \mathbb{V}} M^{\otimes n}\, \mathrm{d}Z_n\right).\end{aligned}$$

Using that (see Lemma B.3)

$$\int \mathbb{1}_{\substack{Z_n \text{ form a} \\ \text{possible cluster}}} M^{\otimes n}\, \mathrm{d}Z_n \leq C_\gamma \mu^{-n+1}\delta^{n-1}, \tag{3.6}$$

$$\int \mathbb{1}_{\|V_{k'}\| \geq \mathbb{V}} M^{\otimes n}\, \mathrm{d}Z_n \leq C_n e^{-\frac{\mathbb{V}^2}{4}} \tag{3.7}$$

we obtain the expected result.

We used that the Gibbs measure is time invariant. □

Hence, if we fix $\delta := \varepsilon^{1/12}$, $\mathbb{V} := |\log\varepsilon|$ and fix $\gamma = 24d$, $\mathbb{P}_\varepsilon(\Upsilon_\varepsilon^c)$ is $O(\varepsilon^d)$.

3.3. **The main part of the cumulant.** We define three kinds of pathology for the pseudotrajectories.

Definition 3.7. *Fix $m \leq n$, collision parameters $(\lambda_1, \cdots, \lambda_l)$ and an initial position $Z_n \in \mathbb{D}^n$.*

- *There is an* overlap *if there are two particles q, q' and a time $\tau \in \delta\mathbb{Z} \cap [0,t]$ such that $|\mathsf{x}_q(\tau) - \mathsf{x}_{q'}(\tau)| \leq \varepsilon$.*

- *Fix a time τ and particles $i_i, \cdots i_k$. We define a graph G^τ with vertex $\{i_1, \cdots i_k\}$, and where (i_a, i_b) is an edge if and only if*

$$\left|\mathsf{x}_{i_a}(\tau, \underline{\lambda}) - \mathsf{x}_{i_b}(\tau, \underline{\lambda})\right| \leq \varepsilon.$$

  There is a multiple encounter *between $i_i, \cdots i_k$ at time τ if G^τ is connected.*
- *Fix $Z_n \in \mathbb{D}^n$ such that there is not a multiple encounters during $[0,t]$.*
  We say that there is a recollision *if the collision graph has a cycle.*

These pathological terms will be considered as error terms, and we forbid them in the main part of our development: we define $\Phi^{0,t}_{m,n}$ as the part of $\Phi^t_{m,n}$ with only non-pathological pseudotrajectories

$$\Phi^{0,t}_{m,n}[h_m](Z_n) := \frac{1}{(n-m)!}\sum_{\mathfrak{l}=1}^{n}\sum_{[m]\subset\lambda_1\subset[n]}\sum_{\substack{(\lambda_2,\cdots,\lambda_{\mathfrak{l}})\\ \in\mathscr{P}^{\mathfrak{l}-1}_{[n]\setminus\lambda_1}}} h_m(\mathsf{Z}_{[m]}(t,\underline{\lambda}))\mathscr{O}_l(Z_{\lambda_1},\cdots,Z_{\lambda_{\mathfrak{l}}}) \times \mathbb{M}^{[m]}_{|\lambda_1|}(Z_{\lambda_1})\prod_{i=2}^{\mathfrak{l}}\mathbb{M}_{|\lambda_i|}(Z_{\lambda_i})\mathbb{1}_{\text{no pathology}} \ .$$

Forgetting the pathological cases allows us to consider a simpler parametrization of the pseudotrajectory. We denote G as the collision graph. The graph G' is constructed by removing the edges $(i,j)_{\tau,s}$ where i and j are in $[m]$. The edges of G' can be ordered: $(i_k, j_k)_{\tau_k, s_k}$ with $\tau_1 < \tau_2 < \cdots < \tau_{n-m}$ (the τ_i are disjoint for almost all initial data).

We can completely reconstruct the pseudotrajectory by considering only the sequence $s_1, \cdots, s_{k-m}$ and the set of tagged particles $[m]$. This is performed in the following definition.

Definition 3.8 (Second definition of a pseudotrajectory)**.** *Fix $m \leq n$, an initial position Z_n and parameters $(s_k)_{k\leq n-m} \in \{\pm 1\}^{n-m}$ and $\omega \subset [n]$ with $|\omega| = m$. In order to construct the pseudotrajectory $\mathsf{Z}(\tau, Z_n, \omega, (s_k)_k)$, we need an auxiliary function $\iota : [0,t] \to \mathbb{N}$, which is increasing, piecewise constant , and left-continuous function.*

At $\tau = 0$, we set $\mathsf{Z}(0, Z_n, \omega, (s_k)_k) := Z_n$ and $\iota(0) := 1$.

Suppose that the pseudo trajectory $\mathsf{Z}(\cdot, Z_n, \omega, (s_k)_k)$ and $\iota(\cdot)$ are constructed in the time interval $[0,\tau]$. At time τ particles i and j meet, i.e.

$$|\mathsf{x}_i(\tau) - \mathsf{x}_j(\tau)| = \varepsilon, \ (\mathsf{x}_i(\tau) - \mathsf{x}_j(\tau))\,(\mathsf{v}_i(\tau) - \mathsf{v}_j(\tau)) > 0.$$

If $(i,j) \in \omega^2$, the two particles interact and we fix $\lim_{\tau'\searrow\tau} \iota(\tau') := \iota(\tau)$.

Otherwise, we fix $\lim_{\tau'\searrow\tau} \iota(\tau') := \iota(\tau) + 1$ and we look at $s_{\iota(\tau)}$. If $s_{\iota(\tau)} = 1$ the two particles interact: as long as $|\mathsf{x}_i - \mathsf{x}_j < \varepsilon$, they follow the dynamic

$$\begin{cases} \dot{\mathsf{x}}_i = \mathsf{v}_i, \ \dot{\mathsf{v}}_i = \dfrac{\alpha}{\varepsilon}\nabla\mathscr{V}\left(\dfrac{\mathsf{x}_i - \mathsf{x}_j}{\varepsilon}\right), \\ \dot{\mathsf{x}}_j = \mathsf{v}_j, \ \dot{\mathsf{v}}_j = \dfrac{-\alpha}{\varepsilon}\nabla\mathscr{V}\left(\dfrac{\mathsf{x}_i - \mathsf{x}_j}{\varepsilon}\right). \end{cases}$$

If $s_{\iota(\tau)} = -1$ the two particles ignore each other: as long as $|\mathsf{x}_i - \mathsf{x}_j| < \varepsilon$,

$$\begin{cases} \dot{\mathsf{x}}_i = \mathsf{v}_i, \ \dot{\mathsf{v}}_i = 0, \\ \dot{\mathsf{x}}_j = \mathsf{v}_j, \ \dot{\mathsf{v}}_j = 0. \end{cases}$$

In both cases, we define $\tau^+ > \tau$ as the first time bigger than τ such that

$$\left|\mathsf{x}_i(\tau^+) - \mathsf{x}_j(\tau^+)\right| = \varepsilon, \ \left(\mathsf{x}_i(\tau^+) - \mathsf{x}_j(\tau^+)\right)\left(\mathsf{v}_i(\tau^+) - \mathsf{v}_j(\tau^+)\right) < 0.$$

We denote $\mathscr{R}^t_{\omega,(s_k)_k} \subset \mathbb{D}^n$ the set of initial parameters such that the pseudotrajectory has a connected collision graph and has neither multiple encounters , nor recollision or nor overlaps. Hence, on $\mathscr{R}^t_{\omega,(s_k)_k} \subset \mathbb{D}^n$, the previous construction has no ambiguity.

We can reconstruct the partition $(\lambda_1, \cdots, \lambda_{\mathfrak{l}})$ for given $(s_i)_{i\leq n-m}$. We define the graph G as a subgraph of the collision graph $\mathscr{G}$ by removing the edges of the form $(i,j)_{\tau,-1}$ (we keep only the interactions). The cluster λ_1 is the union of the connected components in G of the particles $[m]$. The $(\lambda_2, \cdots, \lambda_{\mathfrak{l}})$ are the other connected components.

We have the following equality

$$\Phi^{0,t}_{m,n}[h_m](Z_n) = \frac{1}{(n-m)!} \sum_{(s_k)_{k\le n-m}} \prod_{i=1}^{n-m} s_i \ \mathbb{1}_{\mathscr{R}^t_{[m],(s_k)_k}} h_m(\mathsf{Z}_{[m]}(\tau, Z_n, [m], (s_k)_k)) \tag{3.8}$$

We denote

$$\Phi^{>,t}_{m,n} := \Phi^t_{m,n} - \Phi^{0,t}_{m,n}. \tag{3.9}$$

3.4. **Iteration of the pseudotrajectory development.** The present section is dedicated to the implementation of the short-time sampling (of step δ).

The construction of Section 3.3 is efficient on a short time interval. To raise a long-time result, we need to iterate these kinds of pseudotrajectory representations and to compute sums of the form

$$\sum_{\underline{i}_{n_2}} \Phi^{0,\delta}_{n_1,n_2} \circ \Phi^{0,\delta}_{n_0,n_1}[h_{n_0}](\mathbf{Z}_{\underline{i}_{n_2}}),$$

where $n_0 \le n_1 \le n_2$ are three integers.

Remark 3.5. *In the usual framework, the pseudotrajectories are trees (see for example* [BGSRS23b, BGSRS24, LB25]*): there are more and more particles as we go backward in time. Hence, the development has naturally a semi-group structure, and it is straightforward to continue the development.*

In the present discussion, the pseudotrajectories have a graph structure: particles do not disappear. Hence, we need to work to iterate the process (see Figure 5) .

We introduce the *semi-tree* condition in order to recover some semi-group structure.

Definition 3.9 (semi-tree condition)**.** *Fix $\omega_1 \subset \omega_2$ two finite sets, t and δ two positive real numbers such that $K := \frac{t}{\delta} \in \mathbb{N}^*$. Fix G a collision graph such that*

$$E(G) \subset \left\{ (i,j)_{(\tau,\sigma)} \middle| i,j \in \omega_2 \, ; \, \sigma \in \{\pm 1\} \, ; \, \tau \in (0,t) \right\},$$

and such that G is connected and simply-connected.

We define the decreasing sequence of sets $(\varpi_k)_{k\in[0,K]}$, where ϖ_k is the connected components of ω_1 in the graph of edges

$$\{(i,j) | \exists (\sigma,\tau) \in \{\pm 1\} \times [k\delta, t] \text{ such that } (i,j)_{\tau,\sigma} \in G\}.$$

The graph G verifies the semi-tree condition with respect to ω_1 if for any edges $(i,j)_{\tau,\sigma} \in G$ with τ in $[k\delta,(k+1)\delta[$, one has

$$(i,j) \in \varpi_{k+1} \times (\varpi_k \setminus \varpi_{k+1}) \cup (\varpi_k \setminus \varpi_{k+1}) \times \varpi_{k+1}.$$

Fix $n \in \mathbb{N}^*$, $\omega \subset [n]$ and collision parameters $(s_k)_{k\le n-|\omega|} \in \{\pm 1\}^{n-|\omega|}$. We define $\mathscr{R}^{0,t}_{\omega,(s_k)_k} \subset \mathscr{R}^{0,t}_{\omega,(s_k)_k}$ as the set of initial parameters Z_n such that the collision graph of the pseudotrajectory $\mathsf{Z}(\cdot, Z_n, \omega, (s_k)_k)$ verifies the semi-tree condition with respect to ω.

We need a new definition of pseudotrajectory:

Definition 3.10 (Third definition of pseudotrajectory)**.** *Fix $m \le n$. For a family of parameters $(\omega_1, \omega_2, (s_k)_{k\le n-m})$ with $\omega_1 \subset \omega_2 \subset [n]$, $|\omega_1| = m$ and $(s_k)_{k\le n-m} \in \{\pm 1\}^{n-m}$, we define $\mathrm{Z}(\tau, Z_n, \omega_1, \omega_2, (s_k)_k)$ as*

- *for $\tau \le \delta$,*
$$\mathrm{Z}(\tau, Z_n, \omega_1, \omega_2, (s_k)_{k\le n-m}) := \mathsf{Z}(\tau, Z_n, \omega_2, (s_k)_{k\le n-|\omega_2|}),$$
- *for $\tau > \delta$,*
$$\mathrm{Z}_{\omega_2}(\tau, Z_n, \omega_1, \omega_2, (s_k)_{k\le n-m}) := \mathsf{Z}(\tau-\delta, \mathrm{Z}_{\omega_2}(\delta), \omega_1, (s_k)_{n-|\omega_2|<k\le n-m}),$$
and for all $i \in [n] \setminus \omega_2$
$$\mathrm{z}_i(\tau) := (\mathrm{x}_i(\delta) + (\tau-\delta)\mathrm{v}_i(\delta), \mathrm{v}_i(\delta)).$$

Remark 3.6. *Note that the particles in $[n_2] \setminus \omega_2$ are virtual since time δ: they do not interact with any other particle.*

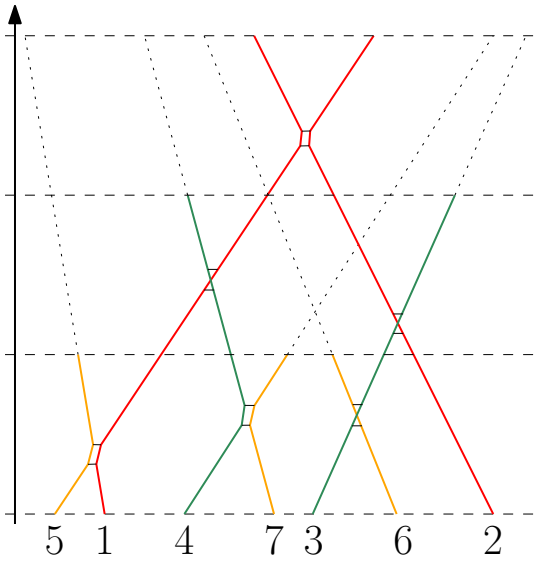

FIGURE 8. Here the pseudotrajectory checks the semi-tree condition, with $\omega_1 = \{1\}$, $\varpi_1 = \{1,2\}$, $\varpi_2 = \{1,2,3,4\}$ and $\varpi_3 = \{1,2,3,4,5,6\}$. In the picture, when the trajectory of a particle is a dotted line, it does not overlap nor interact with any other particle (a crossing with a dotted line is an artifact of the one-dimensional drawing).

Definition 3.11. *We define the collision graph $\mathscr{G}$ of the pseudotrajectory $\mathrm{Z}(\cdot, Z_n, \omega_1, \omega_2, (s_k)_k)$. We define the graph G' as the subgraph of $\mathscr{G}$ with edges*

$$\big\{(i,j)_{\tau,\sigma} \in E(\mathscr{G}),\ \tau \in [\delta, t],\ (i,j) \in \omega_2^2\big\}.$$

Fix $\omega_1 \subset \omega_2 \subset [n]$ and collision parameters $(s_k)_{k \le n - |\omega_1|}$. We define the set

$$\mathscr{R}^t_{\omega_1,\omega_2,(s_k)_k} \subset \left\{ Z_n \in \mathscr{R}^\delta_{\omega_2,(s_k)_{k\le n-|\omega_2|}} \middle| \mathrm{Z}(\cdot, Z_n, \omega_2, (s_k)_{k\le n-|\omega_1|}) \in \mathscr{R}^\delta_{\omega_1,(s_k)_{0\le k-|\omega_2|\le n-|\omega_1|}} \right\}$$

such that G' has no cycle and verifies the semi-tree condition with respect to ω_1.

Fix $m \le n$, ω such that $[m] \subset \omega \subset [n]$ and $(s_k)_{k\le n-m}$ some collision parameters. We split $\mathscr{R}^t_{[m],\omega,(s_k)_k}$ into two pieces: $\mathscr{R}^{>,t}_{[m],\omega,(s_k)_k}$ where the collision graph has at least one cycle and $\mathscr{R}^{0,t}_{[m],\omega,(s_k)_k}$ where the collision graph has no cycle.

For $Z_n \in \mathscr{R}^{0,t}_{[m],\omega,(s_k)_k}$, there are exactly $n-m$ collisions in the collision graph, and ω is not needed to reconstruct the pseudotrajectory. In addition, its collision graph checks the semi-tree conditions. Fixing the parameters $(s_k)_k$, the sets $(\mathscr{R}^{0,t}_{[m],\omega,(s_k)_k})_\omega$ are disjoint, as ω is the union of the connected components of the vertices $1, \cdots, m$ in the graph G'. We deduce that

$$\mathscr{R}^{0,t}_{[m],(s_k)_k} = \bigsqcup_{[m]\subset\omega\subset[n]} \mathscr{R}^{0,t}_{[m],\omega,(s_k)_k}.$$

We introduce now the functionals

$$\Psi^{0,t}_{n_0,n_1}[h_{n_0}] := \frac{1}{(n_1-n_0)!} \sum_{(s_k)_{k\le n_1-n_0}} \prod_{k=1}^{n_1-n_0} s_k\, h_{n_0}\left(\mathrm{Z}_{[n_0]}(t,\cdot,[n_0],(s_k)_k)\right) \mathbb{1}_{\mathscr{R}^{0,t}_{[n_0],(s_k)_k}}, \tag{3.10}$$

$$\Psi^{>,t}_{n_0,n_1}[h_{n_0}] := \frac{1}{(n_1-n_0)!} \sum_{\substack{[n_0]\subset\omega\subset[n_1]\\(s_k)_{k\le n_1-n_0}}} \prod_{k=1}^{n_1-n_0} s_k\, h_{n_0}\left(\mathrm{Z}_{[n_0]}(t,\cdot,[n_0],\omega,(s_k)_k)\right) \mathbb{1}_{\mathscr{R}^{>,t}_{[n_0],\omega,(s_k)_k}}. \tag{3.11}$$

Fix $n_0 \le n_1 \le n_2$ and a test function h_{n_0}. We obtain directly

$$\Phi^{0,\delta}_{n_1,n_2} \circ \Psi^{0,t}_{n_0,n_1}[h_{n_0}](Z_{n_2}) = \frac{1}{(n_2-n_1)!(n_1-n_0)!} \sum_{(s_k)_{k\le n_2-n_0}} \prod_{k=1}^{n_2-n_0} s_k\, h_{n_0}\left(\mathrm{Z}_{[n_0]}(t+\delta, Z_{n_2}, [n_0],[n_1],(s_k)_k)\right) \mathbb{1}_{\mathscr{R}^{t+\delta}_{[n_0],[n_1],(s_k)_k}}.$$

Then, summing on n_1 and ω,

$$\sum_{n_1=n_0}^{n_2} \frac{(n_2-n_1)!(n_1-n_0)!}{(n_2-n_0)!} \sum_{\substack{[n_0]\subset\omega\subset[n_2]\\|\omega|=n_1}} \Phi^{0,\delta}_{n_1,n_2}\circ\Psi^{0,t}_{n_0,n_1}[h_{n_0}](Z_{[n_0]},Z_{\omega\setminus[n_0]},Z_{[n_2]\setminus\omega})$$

$$= \frac{1}{(n_2-n_0)!}\sum_{\substack{[n_0]\subset\omega\subset[n_2]\\(s_k)_{k\le n_2-n_0}}} \prod_{k=1}^{n_2-n_0} s_k\, h_{n_0}\left(\mathrm{Z}_{[n_0]}(t+\delta,Z_{n_2},[n_0],\omega,(s_k)_k)\right)\mathbb{1}_{\mathscr{R}^{t+\delta}_{[n_0],\omega,(s_k)_k}}$$

$$= \Psi^{0,t+\delta}_{n_0,n_2}[h_{n_0}](Z_{n_2}) + \Psi^{>,t+\delta}_{n_0,n_2}[h_{n_0}](Z_{n_2}).$$

Using the symmetry of the summation set,

$$\sum_{n_1=n_0}^{n_2}\sum_{\underline{i}_{n_2}} \Phi^{0,\delta}_{n_1,n_2}\circ\Psi^{0,t}_{n_0,n_1}[h_{n_0}](\mathbf{Z}_{\underline{i}_{n_2}})$$

$$= \sum_{n_1=n_0}^{n_2}\sum_{\underline{i}_{n_2}} \binom{n_2-n_0}{n_1-n_0}^{-1} \sum_{\substack{[n_0]\subset\omega\subset[n_2]\\|\omega|=n_1}} \Phi^{0,\delta}_{n_1,n_2}\circ\Psi^{0,t}_{n_0,n_1}[h_{n_0}](\mathbf{Z}_{\underline{i}_{[n_0]}},\mathbf{Z}_{\underline{i}_{\omega\setminus[n_0]}},\mathbf{Z}_{\underline{i}_{[n_2]\setminus\omega}})$$

$$= \sum_{\underline{i}_{n_2}} \Psi^{0,t+\delta}_{n_0,n_2}[h_{n_0}](\mathbf{Z}_{\underline{i}_{n_2}}) + \sum_{\underline{i}_{n_2}} \Psi^{>,t+\delta}_{n_0,n_2}[h_{n_0}](\mathbf{Z}_{\underline{i}_{n_2}}). \tag{3.12}$$

The functional $\Psi^{0,t}_{m,n}$ are introduced to implement the sampling: for $t>2\delta$ and $\mathbf{Z}_{\mathscr{N}}\in\Upsilon_\varepsilon$ (note that between the first and second lines, we use that by definition $\Phi^{0,\delta}_{1,n}=\Psi^{0,\delta}_{1,n}$)

$$\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t)) = \sum_{n\ge1}\sum_{\underline{i}_n} \Phi^{0,\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(t-\delta)) + \sum_{n\ge1}\sum_{\underline{i}_n} \Phi^{>,\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(t-\delta))$$

$$= \sum_{n'\ge n\ge0}\sum_{\underline{i}_{n'}} \left(\Phi^{0,\delta}_{n,n'}\Psi^{0,\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_{n'}}(t-2\delta)) + \Phi^{>,\delta}_{n,n'}\Psi^{0,\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_{n'}}(t-k\delta))\right)$$

$$+ \sum_{n\ge1}\sum_{\underline{i}_n} \Phi^{>,\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(t-\delta)).$$

This can be rewritten as

$$\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t)) = \sum_{n\ge1}\sum_{\underline{i}_n} \Psi^{0,t}_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(0)) + \sum_{k=1}^{2}\sum_{n\ge1}\sum_{\underline{i}_n} \Psi^{>,k\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(t-k\delta))$$

$$+ \sum_{k=1}^{2}\sum_{1\le n\le n'}\sum_{\underline{i}_{n'}} \Phi^{>,\delta}_{n,n'}\circ\Psi^{0,(k-1)\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_{n'}}(t-k\delta)).$$

The preceding computation can be iterated: for some time t, $\theta<t$ and δ such that $\theta/\delta=K\in\mathbb{N}$, and any initial data $\mathbf{Z}_{\mathscr{N}}\in\Upsilon_\varepsilon$

$$\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t)) = \sum_{n\ge1}\sum_{\underline{i}_n} \Psi^{0,\theta}_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(t-\theta)) + \sum_{k=1}^{K}\sum_{n\ge1}\sum_{\underline{i}_n} \Psi^{>,k\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_n}(t-k\delta))$$

$$+ \sum_{k=1}^{K}\sum_{1\le n\le n'}\sum_{\underline{i}_{n'}} \Phi^{>,\delta}_{n,n'}\circ\Psi^{0,(k-1)\delta}_{1,n}[h](\mathbf{Z}_{\underline{i}_{n'}}(t-k\delta)). \tag{3.13}$$

3.5. **The decomposition of the covariance.** The final ingredient is a second sampling on a longer time scale $\theta\simeq 1/\beta\log|\log\varepsilon|$, which controls the growth of the number of collisions.

Definition 3.12 (Number of particles at time τ)**.** *Fix t and δ such that $t/\delta=K\in\mathbb{N}^*$, parameters $(\{1\},\omega_2,(s_k)_{k\le n-m})$ and admissible initial data $Z_{n_2}\in\mathscr{R}^{0,t}_{\{1\},\omega_2,(s_k)_k}$. We denote $\mathscr{G}$ the collision graph. For $\tau=k\delta$, the* number of particles at time τ, $\mathfrak{n}(\tau)$, *is defined as the size of the connected component of $\{1\}$ in the graph with edges*

$$\{(i,j)|\exists(\sigma,\tau)\in\{\pm1\}\times[k\delta,t] \text{ such that } (i,j)_{\tau,\sigma}\in\mathscr{G}\}.$$

We want to control the number of particles $(\mathfrak{n}(t-k\theta))_k$ such that it grows slower than the geometric sequence 2^k.

Fix $1 \le n_1 \le \cdots \le n_l$. We denote $\underline{n}_l := (n_i)_{i\le l}$. For $t \in ((l-1)\theta, l\theta]$

$$\Psi^{0,t}_{\underline{n}_l}[h] := \frac{1}{(n_l-1)!} \sum_{(s_k)_{k\le n_l-1}} \prod_{k=1}^{n_l-1} s_k h(\mathrm{z}_1(t,\cdot,\{1\},(s_k)_k)) \mathbb{1}_{\mathscr{R}^{0,t}_{\{1\},(s_k)_k}} \prod_{i=1}^{\lfloor t/\theta\rfloor} \mathbb{1}_{\mathfrak{n}(t-i\theta)=n_i}, \tag{3.14}$$

and for $t \in [(l-2)\theta, (l-1)\theta]$

$$\Psi^{>,t}_{\underline{n}_l}[h] := \frac{1}{(n_l-1)!} \sum_{\substack{1\in\omega\subset[n_l]\\(s_k)_{k\le n_l-1}}} \prod_{k=1}^{n_l-1} s_k\, h(\mathrm{z}_1(t,\cdot,\{1\},\omega,(s_k)_k)) \mathbb{1}_{\mathscr{R}^{>,t}_{\{1\},\omega,(s_k)_k}} \prod_{i=1}^{l-2} \mathbb{1}_{\mathfrak{n}(t-i\theta)=n_i}. \tag{3.15}$$

We can iterate the preceding decomposition of $\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t))$. The decomposition is performed until reaching the time 0: denoting $K := t/\theta \in \mathbb{N}$, $K' := \theta/\delta \in \mathbb{N}$, $\theta \simeq 1/\beta \log|\log\varepsilon|$ (for β small enough), and $\delta \simeq \varepsilon^{1/12}$, for almost any initial data $\mathbf{Z}_{\mathscr{N}}(0) \in \mathscr{D}$,

$$\sum_{i=1}^{\mathscr{N}} h(\mathbf{z}_i(t)) = \sum_{\substack{(n_j)_{j\le K}\\0\le n_j-n_{j-1}\le 2^j}} \sum_{\underline{i}_{n_k}} \Psi^{0,t}_{\underline{n}_k}[h]\left(\mathbf{Z}_{\underline{i}_{n_k}}(t)\right) \tag{3.16}$$

$$+ \sum_{1\le k\le K} \sum_{\substack{(n_j)_{j\le k-1}\\0\le n_j-n_{j-1}\le 2^j}} \sum_{n_k\ge 2^k+n_{k-1}} \sum_{\underline{i}_{n_k}} \Psi^{0,k\theta}_{\underline{n}_k}[h]\left(\mathbf{Z}_{\underline{i}_{n_k}}(t-k\theta)\right) \tag{3.17}$$

$$+ \sum_{\substack{0\le k\le K-1\\1\le k'\le K'}} \sum_{\substack{(n_j)_{j\le k}\\0\le n_j-n_{j-1}\le 2^j}} \sum_{n_{k+2}\ge n_{k+1}\ge n_k} \sum_{\underline{i}_{n_k}} \Psi^{>,t-t_s}_{\underline{n}_{k+1}}[h]\left(\mathbf{Z}_{\underline{i}_{n_k}}(t_s)\right) \tag{3.18}$$

$$+ \sum_{\substack{0\le k\le K-1\\1\le k'\le K'}} \sum_{\substack{(n_j)_{j\le k}\\0\le n_j-n_{j-1}\le 2^j}} \sum_{\substack{n_{k+1}\ge n_k\\n_{k+2}\ge n_{k+1}}} \sum_{\underline{i}_{n_{k+2}}} \Phi^{>,\delta}_{n_{k+1},n_{k+2}} \Psi^{0,t_s-\delta}_{\underline{n}_{k+1}}[h]\left(\mathbf{Z}_{\underline{i}_{n_{k+2}}}(t_s)\right) \tag{3.19}$$

where $t_s := t - k\theta - k'\delta$.

Finally, the covariance is split into five parts

$$\mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(h)\zeta^0(g)\right] = G^{\mathrm{main}}_\varepsilon(t) + G^{\mathrm{clus}}_\varepsilon(t) + G^{\mathrm{exp}}_\varepsilon(t) + G^{\mathrm{rec},1}_\varepsilon(t) + G^{\mathrm{rec},2}_\varepsilon(t), \tag{3.20}$$

where we have separated

- the main part,

$$G^{\mathrm{main}}_\varepsilon(t) := \sum_{\substack{(n_j)_{j\le K}\\0\le n_j-n_{j-1}\le 2^j}} \mathbb{E}_\varepsilon\left[\frac{1}{\sqrt{\mu}} \sum_{\underline{i}_{n_K}} \Psi^{0,t}_{\underline{n}_K}[h]\Big(\mathbf{Z}_{\underline{i}_{n_K}}(0)\Big)\zeta^0_\varepsilon(g)\right], \tag{3.21}$$

- the first error due to the symmetric conditioning and the suppression of the overlaps,

$$G^{\mathrm{clus}}_\varepsilon(t) := \mathbb{E}_\varepsilon\Big[\zeta^t_\varepsilon(h)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon^c_\varepsilon}\Big] - \sum_{\substack{(n_j)_{j\le K}\\0\le n_j-n_{j-1}\le 2^j}} \mathbb{E}_\varepsilon\left[\frac{1}{\sqrt{\mu}} \sum_{\underline{i}_{n_K}} \Psi^{0,t}_{\underline{n}_K}[h]\left(\mathbf{Z}_{\underline{i}_{n_K}}(0)\right)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon^c_\varepsilon}\right], \tag{3.22}$$

- the part controlling the growth of the number of particles,

$$G^{\mathrm{exp}}_\varepsilon(t) := \mathbb{E}_\varepsilon\left[(3.17)\times\frac{1}{\sqrt{\mu}}\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right], \tag{3.23}$$

- the part corresponding to non-local recollision,

$$G^{\mathrm{rec},1}_\varepsilon(t) := \mathbb{E}_\varepsilon\left[(3.18)\times\frac{1}{\sqrt{\mu}}\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right], \tag{3.24}$$

- and the part corresponding to local recollision,

$$G^{\mathrm{rec},2}_\varepsilon(t) := \mathbb{E}_\varepsilon\left[(3.19)\times\frac{1}{\sqrt{\mu}}\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]. \tag{3.25}$$

The parts $G_\varepsilon^{\rm clus}(t)$ and $G_\varepsilon^{\rm exp}(t)$ are estimated by (5.1):

$$|G_\varepsilon^{\rm exp}(t)+G_\varepsilon^{\rm rec}(t)|\leq C\|g\|_0\|h\|_0\left(\varepsilon^{1/3}(C\tfrac{t}{\mathfrak{d}})^{2^K}+\tfrac{\theta t^2}{\mathfrak{d}^3}\right),$$

the part $G_\varepsilon^{\rm rec,1}(t)$ is estimated by (7.1): for some $\mathfrak{a}>0$ depending only on the dimension,

$$\left|G_\varepsilon^{\rm rec,1}(t)\right|\leq\|g\|_0\|h\|_0(Ct)^{2^K+d+9}\varepsilon^{\mathfrak{a}/2},$$

the part $G_\varepsilon^{\rm rec,2}(t)$ is bounded at (8.1):

$$\left|G_\varepsilon^{\rm rec,2}(t)\right|\leq C\|h\|_0\|g\|_0(C\tfrac{t}{\mathfrak{d}})^{2^{K+1}}\varepsilon^{\frac{\mathfrak{a}}{2}},$$

and the convergence of $G_\varepsilon^{\rm main}(t)$ is given by (6.19):

$$G_\varepsilon^{\rm main}(t)=\int_{\mathbb{D}}h(z)\mathbf{g}_\alpha(t,z)M(z)dz+O\left(\left(C\tfrac{\theta t}{\mathfrak{d}^2}+\varepsilon^{\mathfrak{a}}K2^{K^2}(\tfrac{Ct}{\mathfrak{d}})^{2^{K+1}}\right)\|h\|_1\|g\|_1\right)$$

where $\mathbf{g}_\alpha(t,z)$ is the solution of the linearized Boltzmann equation (2.16). Combining these four estimations and that $K\leq T/\theta$, we obtain the expected bound (2.15)

$$\mathbb{E}_\varepsilon\left[\zeta_\varepsilon^t(g)\zeta_\varepsilon^0(h)\right]=\int_{\mathbb{D}}h(z)\mathbf{g}_\alpha(t,z)M(z)dz+O\left(\left(C\tfrac{\theta T^2}{\mathfrak{d}^3}+\varepsilon^{\mathfrak{a}}(\tfrac{CT}{\mathfrak{d}})^{2^{\frac{T}{\theta}+1}}\right)\|h\|_1\|g\|_1\right)\tag{3.26}$$

Remark 3.7. *In this section, we have defined three different pseudotrajectories:*

- *in Definition 3.3 we have defined the general definition of pseudotrajectory, which is used in the estimation of pathological recollision $G_\varepsilon^{rec,1}(t)$,*
- *the pseudotrajectories of Definition 3.8 have no recollision and are used to treat $G_\varepsilon^{main}(t)$, $G_\varepsilon^{clust}(t)$ and $G_\varepsilon^{exp}(t)$,*
- *Definition 3.10 describes pseudotrajectories with non-pathological recollision. They are used to bound $G_\varepsilon^{rec,2}(t)$.*

Remark 3.8 (Comparison with the strategy of [LB25, LB23])**.** *As explained in Remark 3.5, we used in* [LB25] *a* tree pseudotrajectory *representation, while we use in the present article the graph pseudotrajectories. There are three advantages to this construction:*

- *The tree pseudotrajectories have already been used in the description of gas interacting via compactly supported potential* [Gra49, Kin75, GSRT13, PSS14, Ayi17, Cat18]*. One of the encountered difficulties is the treatment of multiple interactions, which have to be treated separately from binary interactions even before imposing some conditioning. Our presentation avoids this difficulty.*
- *In* [LB25]*, we needed two conditionings (see Section 2.3 of* [LB25]*). The first one is symmetric, and looks like Υ_ε (we denote it $\tilde\Upsilon_\varepsilon$) : we impose the same condition on all the particles. We also needed a second conditioning, which is asymmetric: we construct an indicator function $\mathscr{X}_{\underline{i}_n}:\mathscr{D}\to\{0,1\}$ to forbid a recollision on a small slice $[k\delta,(k+1)\delta]$ that involves one of the particles of $\underline{i}_n$. Then we made the decomposition*

$$\sum_{\underline{i}_n}\Psi_n(\mathbf{Z}_{\underline{i}_n})=\sum_{\underline{i}_n}\Psi_n(\mathbf{Z}_{\underline{i}_n})\mathscr{X}_{\underline{i}_n}(\mathbf{Z}_{\mathscr{N}})+\sum_{\underline{i}_n}\Psi_n(\mathbf{Z}_{\underline{i}_n})(1-\mathscr{X}_{\underline{i}_n}(\mathbf{Z}_{\mathscr{N}})).$$

  The use of the graph pseudotrajectories allows us to avoid this second conditioning.
- *In* [LB25]*, we used the following conditioning: let $Z_n\in\mathbb{D}^n$ be a particle configuration. We consider the graph G of vertices $\{1,\cdots,n\}$ and of edges*

$$\{(i,j)\in[1,m]^2,\ d(x_i,x_j)<2\delta\mathbb{D}\}.$$

  The coordinates Z_n form a distance cluster *if the graph G is connected.*

  We construct the conditioning $\tilde\Upsilon_\varepsilon$ as we have constructed Υ_ε, by replacing the notion of possible cluster *by the distance clusters (we suppose that there is not a distance cluster larger than some integer γ).*

  The conditioning $\tilde\Upsilon_\varepsilon$ is more constraining than Υ_ε: we obtained in Proposition 2.3 of [LB25]

$$\mathbb{P}_\varepsilon\left(\tilde\Upsilon_\varepsilon^c\right)\leq C_\gamma\frac{t}{\delta}\left(\mu_\varepsilon\delta^{-1}\left(\mu_\varepsilon\delta^d\mathbb{V}^d\right)^\gamma+\mu_\varepsilon^\gamma e^{-\mathbb{V}^2/4}\right),$$

  which is much larger than our estimation on $\mathbb{P}_\varepsilon\left(\tilde\Upsilon_\varepsilon^c\right)$.

  The conditioning $\tilde\Upsilon_\varepsilon$ is only useful if we take δ small enough (of order $\varepsilon^{1-\frac{1}{2d}}$ for example). The problem is that the gain of smallness due to recollisions has to be smaller than δ. If we restrict ourselves to the hard sphere setting, there is no problem, as one can obtain a gain of

order $\varepsilon|\log\varepsilon|$ (see the Appendix of [BGSRS23b]*). In Assumption 2.3.1 of* [LB23]*, we proposed a family of interaction potentials for which the same estimation holds.*

The cluster decomposition of this paper allows us to take a much larger $\delta = \varepsilon^{\frac{1}{12}}$ and then use the strategy of [PSS14] *to bound the recollisions. This improvement is possible because we used the graph pseudo-trajectory.*

4. Quasi-orthogonality estimates

The different error terms obtained in the previous section are of the form

$$\mathbb{E}_\varepsilon\left[\sum_{\underline{i}_n} \Phi_n[h](\mathbf{Z}_{\underline{i}_n}(t_s))\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]$$

where the $\Phi_n : L^\infty(\mathbb{D}) \to L^\infty(\mathbb{D}^n)$ are continuous functionals. In order to bound the errors, we need an $L^2(\mathbb{P}_\varepsilon)$ bound of

$$\hat{\Phi}_n = \frac{1}{\mu^n}\sum_{\underline{i}_n} \Phi_n[h](\mathbf{Z}_{\underline{i}_n}) - \mathbb{E}[\Phi_n].$$

The following section is dedicated to the derivation of such an estimate, using detailed estimations on the functionals $\Phi_n[h]$. We will use, in particular, that we can bound the $\Phi_n[h](Z_n)$ by looking only at the relative positions of particles inside Z_n.

Definition 4.1. *We denote for $y \in \mathbb{T}$ the translation operator*

$$\mathrm{tr}_y : \begin{cases} \mathbb{D}^n & \to & \mathbb{D}^n \\ (X_n, V_n) & \mapsto & (x_1 + y, \cdots, x_n + y, V_n) \end{cases}. \tag{4.1}$$

Fix n, m two integers, $g_n : \mathbb{D}^n \to \mathbb{R}$, $h_m : \mathbb{D}^m \to \mathbb{R}$ two functions, and $l \in [0, \min(n, m)]$. We define the multiplication on l variable $\circledast_l$ as

$$g_n \circledast_l h_m(Z_{n+m-l}) := \frac{1}{(n+m-l)!n!m!} \sum_{\substack{\sigma,\sigma',\sigma''\in\mathfrak{S}([n+m-l]) \\ \sigma'_{|[1,n]^c}=\mathrm{Id} \\ \sigma''_{|[n-l,2n-l]^c}=\mathrm{Id}}} g_n(Z_{\sigma\sigma'([1,n])})h_m(Z_{\sigma\sigma''([n+1-l,n+m-l])}). \tag{4.2}$$

where $\mathfrak{S}(\omega)$ is the group of permutation of ω.

Proposition 4.1. *Fix $m < n$ two positive integers, and $g_n : \mathbb{D}^n \to \mathbb{R}$, $h_m\mathbb{D}^m \to \mathbb{R}$ two functions such that there exists a finite sequence $(c_0, c_0', c_1, \cdots, c_n) \in \mathbb{R}_+^{n+2}$ bounding g_n, h_m in the following way:*

$$\int_{x_1=0} \sup_{y\in\mathbb{T}} \left|g_n\left(\mathrm{tr}_y Z_n\right)\right| \frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{\frac{nd}{2}}}\,\mathrm{d}X_{2,n}\,\mathrm{d}V_n \le c_0, \tag{4.3}$$

$$\int_{x_1=0} \sup_{y\in\mathbb{T}} \left|h_m\left(\mathrm{tr}_y Z_m\right)\right| \frac{e^{-\mathscr{H}_m(Z_m)}}{(2\pi)^{\frac{md}{2}}}\,\mathrm{d}X_{2,m}\,\mathrm{d}V_m \le c_0' \tag{4.4}$$

and for all $l \in [1, m]$

$$\int_{x_1=0} \sup_{y\in\mathbb{T}} \left|g_n \circledast_l h_m\left(\mathrm{tr}_y Z_{n+m-l}\right)\right| \frac{e^{-\mathscr{H}_{n+m-l}(Z_{n+m-l})}}{(2\pi)^{\frac{(n+m-l)d}{2}}}\,\mathrm{d}X_{2,n+m-l}\,\mathrm{d}V_{n+m-l} \le \frac{\mu^{l-1}}{n^l}c_l. \tag{4.5}$$

There exists a constant $C > 0$ depending only on the dimension such that

$$\mathbb{E}_\varepsilon\left[g_n\right| = \int g_n(Z_n)\frac{e^{-\mathscr{H}(Z_n)}}{(2\pi)^{nd/2}}\,\mathrm{d}Z_n + O(C^n c_0 \tfrac{\varepsilon}{\mathfrak{d}}) \tag{4.6}$$

and

$$\mathbb{E}_\varepsilon\left[\mu\hat{g}_n\hat{h}_m\right] = \sum_{l=1}^m \binom{n}{l}\binom{m}{l}\frac{l!}{\mu^{l-1}}\int g_n \circledast_l h_m(Z_{n+m-l})\frac{e^{-\mathscr{H}_{n+m-l}(Z_{n+m-l})}}{(2\pi)^{\frac{(n+m-l)d}{2}}}\,\mathrm{d}Z_{n+m-l} + O\big(\big(c_0c_0' + \max_{1\le l\le m} c_l\big)\,C^{n+m}\tfrac{\varepsilon}{\mathfrak{d}}\big). \tag{4.7}$$

In particular

$$\left|\mathbb{E}_\varepsilon\left[g_n\right]\right| = O(C^n c_0), \quad \left|\mathbb{E}_\varepsilon\left[\mu \hat{g}_n \hat{h}_m\right]\right| \leq C^{n+m}\left(\max_{1\leq l\leq m} c_l + c_0 c_0' \frac{\varepsilon}{\mathfrak{d}}\right). \tag{4.8}$$

Remark 4.1. *This theorem is the counterpart of Proposition 3.1 of* [BGSRS23b].

Proof of Proposition 4.1.

- We begin with the proof of (4.6).

Using invariance of $\mathscr{H}_p$ under permutation,

$$\begin{aligned}
\mathbb{E}_\varepsilon[g_n] &= \frac{1}{\mathscr{Z}\mu^n}\sum_{p\geq n}\frac{\mu^p}{p!}\int \sum_{\substack{(i_1,\cdots i_n)\\ \forall k, i_k\leq p}} g_n(Z_{\underline{i}_n}) e^{-\mathscr{H}_p(Z_p)}\frac{\mathrm{d}Z_p}{(2\pi)^{dp/2}} \\
&= \frac{1}{\mathscr{Z}\mu^n}\sum_{p\geq n}\frac{\mu^p}{p!}\frac{p!}{(n-p!)}\int g_n(Z_n) e^{-\mathscr{H}_p(Z_p)}\frac{\mathrm{d}Z_p}{(2\pi)^{dp/2}} \\
&= \frac{1}{\mathscr{Z}}\sum_{p\geq 0}\frac{\mu^p}{p!}\int g_n(Z_n) e^{-\mathscr{V}_{n+p}(X_n,\underline{X}_p)} M^{\otimes n}\,\mathrm{d}Z_n\,\mathrm{d}\underline{X}_p.
\end{aligned}$$

We recall the notation

$$\mathscr{V}_n(X_n) := \alpha \sum_{1\leq i<j\leq n} \mathscr{V}\left(\frac{x_i - x_j}{\varepsilon}\right)$$

and we denote in the following $\Omega := \{X_n, \underline{x}_1, \cdots, \underline{x}_p\}$ and for $X, Y \in \Omega$,

$$\varphi(\underline{x}_i, \underline{x}_j) := \exp\left(-\alpha\mathscr{V}\left(\frac{\underline{x}_i - \underline{x}_j}{\varepsilon}\right)\right) - 1,\ \varphi(X_N, \underline{x}_j) := \exp\left(-\alpha\sum_{i=1}^N \mathscr{V}\left(\frac{x_i - \underline{x}_j}{\varepsilon}\right)\right) - 1. \tag{4.9}$$

Defining $d((x_1,\cdots,x_n),(y_1,\cdots,y_m))$ as the minimum of the $|x_i - y_j|$, we can bound φ by

$$-\mathbb{1}_{d(X,Y)<\varepsilon} \leq \varphi(X,Y) \leq 0.$$

We decompose $\exp\left(-\mathscr{V}_{n+p}(X_{n+1}, \underline{X}_p)\right)$

$$e^{-\mathscr{V}_{n+p}(X_{n+1},\underline{X}_p)} = e^{-\mathscr{V}_n^\varepsilon(X_n)}\prod_{\substack{(X,Y)\in\Omega^2\\ X\neq Y}}(1+\varphi(X,Y)) = e^{-\mathscr{V}_n(X_n)}\sum_{G\in\mathscr{G}(\Omega)}\prod_{(X,Y)\in E(G)}\varphi(X,Y)$$

where $\mathscr{G}(\Omega)$ is the set of non-oriented graphs on Ω and $E(G)$ the set of edges of G. Denoting by $\mathscr{C}(\omega)$ the set of connected graphs on ω,

$$\begin{aligned}
&\exp\left(-\mathscr{V}_{n+p}(X_n, \underline{X}_p)\right) \\
&= \sum_{\omega\subset[1,p]}\left(e^{-\mathscr{V}_n(X_n)}\sum_{G\in\mathscr{C}(\omega\cup\{X_n\})}\prod_{(X,Y)\in E(G)}\varphi(X,Y)\sum_{G\in\mathscr{G}([p]\setminus\omega)}\prod_{(X,Y)\in E(G)}\varphi(X,Y)\right) \\
&= e^{-\mathscr{V}_n(X_n)}\sum_{\omega\subset[1,p]}\left(e^{-\mathscr{V}_{p-|\omega|}(\underline{X}_{[p]\setminus\omega})}\sum_{G\in\mathscr{C}(\omega\cup\{X_n\})}\prod_{(X,Y)\in E(G)}\varphi(X,Y)\right) \\
&=: e^{-\mathscr{V}_n(X_n)}\sum_{\omega\subset[1,p]} e^{-\mathscr{V}_{p-|\omega|}(\underline{X}_{[p]\setminus\omega})}\psi^n_{|\omega|}(X_n, \underline{X}_\omega)\ .
\end{aligned} \tag{4.10}$$

Thus, using exchangeability, $\mathbb{E}_\varepsilon[g_n]$ is equal to

$$\begin{aligned}
&\frac{1}{\mathscr{Z}}\sum_{p\geq 0}\sum_{p_1+p_2=p}\frac{\mu^p}{p!}\frac{p!}{p_1!p_2!}\int g_n(Z_n)\psi^n_{p_1}(X_n,\underline{X}_{p_1}) e^{-\mathscr{V}_{p_2}(\underline{X}'_{p_2})}\frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{\frac{nd}{2}}}\mathrm{d}Z_n\mathrm{d}\underline{X}_{p_1}\mathrm{d}\underline{X}'_{p_2} \\
&= \left(\frac{1}{\mathscr{Z}}\sum_{p\geq 0}\frac{\mu^p}{p!}\int e^{-\mathscr{V}_p(\underline{X}_p)}\,\mathrm{d}\underline{X}_p\right)\left(\sum_{p\geq 0}\frac{\mu^p}{p!}\int g_n(Z_n)\psi^n_p(X_n,\underline{X}_p)\frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{\frac{nd}{2}}}\,\mathrm{d}Z_n\,\mathrm{d}\underline{X}_p\right) \\
&= \sum_{p\geq 0}\frac{\mu^p}{p!}\int g_n(Z_n)\psi^n_p(X_n,\underline{X}_p)\frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{\frac{nd}{2}}}\,\mathrm{d}Z_n\,\mathrm{d}\underline{X}_p.
\end{aligned} \tag{4.11}$$

We recall Penrose's tree inequality (see [Pen63, BGSRS23a, Jan]), for function φ defined in (4.9),

$$\left|\sum_{C\in\mathscr{C}(\Omega)}\prod_{(X,Y)\in E(C)}\varphi(X,Y)\right| \le \sum_{T\in\mathscr{T}(\Omega)}\prod_{(X,Y)\in E(T)}|\varphi(X,Y)| \le \sum_{T\in\mathscr{T}(\Omega)}\prod_{(X,Y)\in E(T)}\mathbb{1}_{d(X,Y)<\varepsilon} \tag{4.12}$$

with $\mathscr{T}(\Omega)$ the set of trees (minimally connected graphs) on Ω. Fix $\mathrm{tr}_{-x_1}X_n$ (the relative position between particles $1,\cdots,n$). Integrating a constraint $\varphi(\underline{x}_i,\underline{x}_j)$ provides a factor $\mathbf{c}_d\varepsilon^d$, $\varphi(X_n,\underline{x}_j)$ a factor $n\mathbf{c}_d\varepsilon^d$ (where $\mathbf{c}_d$ is the volume of a sphere of diameter 1). As there are (see for example the Section 2 of [BGSRS23a] or [Jan])

$$\frac{(p-1)!}{(d_0-1)!(d_1-1)!\cdots(d_p-1)!}$$

trees with specified vertex degrees $d_0,\cdots,d_p$ associated to vertices X_n, $\underline{x}_1,\cdots,\underline{x}_p$, we get

$$\begin{aligned}\left|\int\psi_p^n(X_n\underline{X}_p)\,\mathrm{d}\underline{X}_p\right| &\le \sum_{\substack{d_1,\cdots,d_p\ge1\\ d_0+\cdots+d_p=2p}}\frac{(p-1)!}{(d_0-1)!(d_1-1)!\cdots(d_p-1)!}n^{d_0}(\mathbf{c_d}\varepsilon^d)^p\\ &\le (p-1)!(\mathbf{c_d}\varepsilon^d)^p\left(\sum_{d_0\ge1}\frac{n^{d_0}}{(d_0-1)!}\right)\left(\sum_{d_1\ge1}\frac{1}{(d_1-1)!}\right)\cdots\left(\sum_{d_p\ge1}\frac{1}{(d_p-1)!}\right)\\ &\le (p-1)!\,ne^n\left(e\mathbf{c_d}\varepsilon^d\right)^p.\end{aligned} \tag{4.13}$$

We can integrate on the rest of the parameters using (4.3). Hence, as $\psi_0^n(Z_n)=\frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{nd/2}}$

$$\begin{aligned}\left|\mathbb{E}_\varepsilon[g_n]|-\int g_n(Z_n)\frac{e^{-\mathscr{H}(Z_n)}}{(2\pi)^{nd/2}}\,\mathrm{d}Z_n\right| &\le \sum_{p\ge1}\frac{(p-1)!\,ne^n\left(e\mathbf{c_d}\mu\varepsilon^d\right)^p}{p!}\int|g_n(Z_n)|\frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{\frac{nd}{2}}}\,\mathrm{d}Z_n\\ &\le \frac{c_0C^n\varepsilon}{\mathfrak{d}}\sum_{p\ge0}(C\varepsilon/\mathfrak{d})^p.\end{aligned}$$

For some constant C depending only on the dimension. The series converges for ε small enough as $\mathfrak{d}\gg\varepsilon$. This concludes the proof of (4.6).

- We treat now (4.7). Recall first that

$$\mathbb{E}_\varepsilon\left[\mu\hat{g}_n\hat{h}_m\right]=\frac{1}{\mu^{n+m-1}}\mathbb{E}_\varepsilon\left[\sum_{\underline{i}_n}g_n(\mathbf{Z}_{\underline{i}_n})\sum_{\underline{j}_m}h_m(\mathbf{Z}_{\underline{j}_m})\right]-\mu\mathbb{E}_\varepsilon\left[g_n\right]\mathbb{E}_\varepsilon\left[h_m\right].$$

Let us count the number of ways such that $\underline{i}_n$ and $\underline{j}_m$ can intersect on a set of length l. We have to choose two sets $A\subset[n]$ and $A'\subset[m]$ of length l, and a bijection $\sigma:A\to A'$ such that for all indices $k\in A$, $i_k=j_{\sigma k}$ and that $\underline{i}_{A^c}$ does not intersect $\underline{j}_{(A^c)'}$. Thus, using the symmetry,

$$\begin{aligned}\mathbb{E}_\varepsilon\left[\mu\,\hat{g}_n\hat{h}_m\right]=&\sum_{l=1}^m\binom{n}{l}\binom{m}{l}\frac{l!}{\mu^{l-1}}\mathbb{E}_\varepsilon\left[g_n\circledast_l h_m\right]\\ &+\mu\left(\mathbb{E}_\varepsilon\left[\frac{1}{\mu^{n+m}}\sum_{\underline{i}_{n+m}}g_n(\mathbf{Z}_{\underline{i}_n})h_m(\mathbf{Z}_{\underline{i}_{n+1,n+m}})\right]-\mathbb{E}_\varepsilon\left[g_n\right]\mathbb{E}_\varepsilon\left[h_m\right]\right).\end{aligned}$$

To estimate the error term in (4.7), we need to compute

$$\begin{aligned}&\mathbb{E}_\varepsilon\left[\frac{1}{\mu^{n+m}}\sum_{\underline{i}_{n+m}}g_n(\mathbf{Z}_{\underline{i}_n})h_m(\mathbf{Z}_{\underline{i}_{n+1,n+m}})\right]\\ &=\frac{1}{\mathscr{Z}}\sum_{p\ge0}\frac{\mu^p}{p!}\int g_n(Z_n)h_m(Z_m')\exp\left(-\mathscr{V}_{n+m+p}(X_n,X_m',\underline{X}_p)\right)M^{\otimes n}\,\mathrm{d}Z_nM^{\otimes m}\,\mathrm{d}Z_m'\,\mathrm{d}\underline{X}_p.\end{aligned}$$

We denote in the following $\Omega := \{X_n, X'_m, \underline{x}_1, \cdots, \underline{x}_p\}$, and we decompose

$$\begin{aligned}\exp\left(-\mathscr{V}_{n+m+p}(X_n, X'_m, \underline{X}_p)\right) &= e^{-\mathscr{V}_n(X_n)} e^{-\mathscr{V}_m(X'_m)} \prod_{\substack{(X,Y)\in\Omega^2 \\ X\neq Y}} (1+\varphi(X,Y)) \\ &= e^{-\mathscr{V}_n(X_n)} e^{-\mathscr{V}_m(X'_m)} \sum_{G\in\mathscr{G}(\Omega)} \prod_{(X,Y)\in E(G)} \varphi(X,Y)\end{aligned}$$

where

$$\varphi(X_n, X'_m) := \exp\left(-\alpha \sum_{i=1}^{n}\sum_{j=1}^{m} \mathscr{V}\left(\frac{x_i - x'_j}{\varepsilon}\right)\right).$$

We make a partition depending on the connected components of X_n and X'_m in G,

$$\begin{aligned}\exp\left(-\mathscr{V}_{n+m+p}(X_n, X'_m, \underline{X}_p) + \mathscr{V}_n(X_n) + \mathscr{V}_m(X'_m)\right)\right) &= \sum_{\omega\subset[1,p]} \psi^{n,m}_{|\omega|}(X_n, X'_m, \underline{X}_\omega) e^{-\mathscr{V}_{p-|\omega|}(\underline{X}_{[p]\setminus\omega})} \\ &+ \sum_{\substack{\omega_1,\omega_1\subset[1,p] \\ \omega_1\cap\omega_2=\emptyset}} \psi^{n}_{|\omega_1|}(X_n, \underline{X}_{\omega_1}) \psi^{m}_{|\omega_2|}(X'_m, \underline{X}_{\omega_2}) e^{-\mathscr{V}_{p-|\omega_1\cup\omega_2|}(\underline{X}_{[p]\setminus(\omega_1\cup\omega_2)})}.\end{aligned}$$

where the first line corresponds to X_n and X'_m in the same connected components, and second corresponds to X_n and X'_m in disjoint connected components. In the preceding equation, we denote

$$\psi^{n,m}_{|\omega|}(X_n, X'_m, \underline{X}_\omega) = \sum_{G\in\mathscr{C}(\omega\cup\{X_n, X'_m\})} \prod_{(X,Y)\in E(G)} \varphi(X,Y).$$

Permutating the indices and using (4.11), we obtain the following equality:

$$\begin{aligned}\frac{1}{\mathscr{Z}} \sum_{p\geq 0} \frac{\mu^p}{p!} \int g_n(Z_n) h_m(Z'_m) \sum_{\substack{\omega_1,\omega_1\subset[1,p] \\ \omega_1\cap\omega_2=\emptyset}} \psi^{n}_{|\omega_1|}(X_n, \underline{X}_{\omega_1}) \psi^{m}_{|\omega_2|}(X'_m, \underline{X}_{\omega_2}) e^{-\mathscr{V}_{|(\omega_1\cup\omega_2)^c|}(\underline{X}_{(\omega_1\cup\omega_2)^c})} \\ \times \frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{\frac{nd}{2}}} \mathrm{d}Z_n \frac{e^{-\mathscr{H}_m(Z'_m)}}{(2\pi)^{\frac{md}{2}}} \mathrm{d}Z'_m \, \mathrm{d}\underline{X}_p \\ = \frac{1}{\mathscr{Z}} \sum_{p\geq 0} \sum_{p_1+p_2+p_3=p} \frac{\mu^p}{p!} \frac{p!}{p_1!p_2!p_3!} \int g_n(Z_n) h_{n'}(Z'_{n'}) \psi^{n}_{p_1}(X_n, \underline{X}_{p_1}) \psi^{1}_{p_2}(x_{n+1}, \underline{X}'_{p_2}) \\ \times \Big(\frac{e^{-\mathscr{H}_n(Z_n)}}{(2\pi)^{\frac{nd}{2}}} \mathrm{d}Z_n d\underline{X}_{p_1}\Big) \Big(\frac{e^{-\mathscr{H}_m(Z'_m)}}{(2\pi)^{\frac{md}{2}}} \mathrm{d}Z'_{n'} \, \mathrm{d}\underline{X}'_{p_2}\Big) \Big(e^{-\mathscr{V}_{p_3}(\underline{X}''_{p_3})} \, \mathrm{d}\underline{X}''_{p_3}\Big) \\ = \mathbb{E}_\varepsilon[g_n] \mathbb{E}_\varepsilon[h_{n'}],\end{aligned}$$

and in the same way

$$\begin{aligned}\frac{1}{\mathscr{Z}} \sum_{p\geq 0} \frac{\mu^p}{p!} \int g_n(Z_n) h_m(Z'_m) \sum_{\omega\subset[1,p]} \psi^{n,m}_{|\omega|}(X_n, X'_m, \underline{X}_\omega) e^{-\mathscr{V}_{|\omega^c|}(\underline{X}_{\omega^c})} \\ \times \frac{e^{-\mathscr{H}_n(Z_n) - \mathscr{H}_m(Z'_m)}}{(2\pi)^{\frac{(n+m)d}{2}}} \mathrm{d}Z_n \, \mathrm{d}Z'_m \, \mathrm{d}\underline{X}_p \\ = \frac{1}{\mathscr{Z}} \sum_{p\geq 0} \sum_{p_1+p_2=p} \frac{\mu^p}{p!} \frac{p!}{p_1!p_2!} \int g_n(Z_n) h_m(Z'_m) \psi^{n,m}_{|\omega|}(X_n, X'_m, \underline{X}_{p_1}) e^{-\mathscr{V}_{p_2}(\underline{X}'_{p_2})} \\ \times \frac{e^{-\mathscr{H}_n(Z_n) - \mathscr{H}_m(Z'_m)}}{(2\pi)^{\frac{(n+m)d}{2}}} \mathrm{d}Z_n \, \mathrm{d}Z'_m \, \mathrm{d}\underline{X}_{p_1} \, \mathrm{d}\underline{X}'_{p_2} \\ = \sum_{p_1\geq 0} \frac{\mu^p}{p_1!} \int g_n(Z_n) h_m(Z'_m) \psi^{n,m}_{|\omega|}(X_n, X'_m, \underline{X}_{p_1}) \frac{e^{-\mathscr{H}_n(Z_n) - \mathscr{H}_m(Z'_m)}}{(2\pi)^{\frac{(n+m)d}{2}}} \mathrm{d}Z_n \, \mathrm{d}Z'_m \, \mathrm{d}\underline{X}_{p_1} \, \mathrm{d}\underline{X}'_{p_2}.\end{aligned}$$

We will use again the Penrose tree inequality,

$$\left|\psi^{n,m}_{|\omega|}(X_n, X'_m, \underline{X}_{p_1})\right| \leq \sum_{T\in\mathscr{T}(\Omega)} \prod_{(X,Y)\in E(T)} |\varphi(X,Y)| \leq \sum_{T\in\mathscr{T}(\Omega)} \prod_{(X,Y)\in E(T)} \mathbb{1}_{d(X,Y)<\varepsilon}.$$

First, we fix $\mathrm{tr}_{-x_1} X_n$ and $\mathrm{tr}_{-x'_1} X'_m$. Integrating a constraint $\varphi(\underline{x}_i, \underline{x}_j)$ provides a factor $\mathbf{c}_d\varepsilon^d$, $\varphi(X_n, \underline{x}_j)$ a factor $n\mathbf{c}_d\varepsilon^d$, $\varphi(X'_m, \underline{x}_j)$ a factor $m\mathbf{c}_d\varepsilon^d$, and $\varphi(X_n, X'_m)$ a factor $nm\mathbf{c}_d\varepsilon^d$. Denoting $d_0, d'_0, d_1 \cdots, d_p$

the degrees of X_n, X'_m, $\underline{x}_1, \cdots, \underline{x}_m$ and $\hat{x}_1 := x_1 - x'_1$,

$$\begin{aligned}
&\left| \int \psi^{n,m}_{|\omega|}(X_n, X'_m, \underline{X}_{p_1}) d\underline{X}_p d\hat{x}_1 \right| \\
&\qquad \leq \sum_{\substack{d'_0, d_0, \cdots, d_p \geq 1 \\ d'_0 + d_0 + \cdots + d_p = 2p}} \frac{p!}{(d'_0 - 1)(d_0 - 1)! \cdots (d_p - 1)!} n^{d_0} m^{d'_0} (\mathbf{c_d} \varepsilon^d)^{+1} \\
&\qquad \leq p! \big(\mathbf{c_d} \varepsilon^d\big)^{p+1} nm\, e^{n+m+p}.
\end{aligned} \tag{4.14}$$

We can integrate on the rest of the parameters using (4.3) and (4.4), and finally

$$\begin{aligned}
&\mu \left(\mathbb{E}_\varepsilon \left[\frac{1}{\mu^{n+m}} \sum_{\underline{i}_{n+m}} g_n(\mathbf{Z}_{\underline{i}_n}) h_m(\mathbf{Z}_{\underline{i}_{n+1,n+m}}) \right] - \mathbb{E}_\varepsilon [g_n] \, \mathbb{E}_\varepsilon [g] \right). \\
&\leq c_0 c'_0 \mu \sum_{p \geq 0} \frac{\mu^p}{p!} p! \big(\mathbf{c_d} \varepsilon^d\big)^{p+1} nm\, e^{n+m+p} \\
&\leq \mu \varepsilon^d nm (\mathbf{c_d} e)^{n+m} c_0 c'_0 \sum_{p \geq 0} (e \mathbf{c_d} \varepsilon)^p \\
&\leq (\varepsilon / \mathfrak{d}) C^{n+m+1} \sum_{p \geq 0} (e \mathbf{c_d} \varepsilon / \mathfrak{d})^p
\end{aligned}$$

which converges for ε small enough.

- To conclude the proof, we apply the estimation (4.6) to (4.7):

$$\begin{aligned}
\mathbb{E}_\varepsilon \Big[\mu\, \hat{g}_n \hat{h}_m \Big] &= \sum_{l=1}^m \binom{n}{l} \binom{m}{l} \frac{l!}{\mu^{l-1}} \mathbb{E}_\varepsilon \big[g_n \circledast_l h_m \big] + O(C^{n+m} c_0 c'_0 \tfrac{\varepsilon}{\mathfrak{d}}) \\
&= \sum_{l=1}^m \binom{n}{l} \binom{m}{l} \frac{l!}{\mu^{l-1}} \int g_n \circledast_l h_m(Z_{n+m-l}) \frac{e^{-\mathscr{H}_{n+m-l}(Z_{n+m-l})}}{(2\pi)^{\frac{(n+m-l)d}{2}}} \, \mathrm{d}Z_{n+m-l} \\
&\quad + O\left(\frac{\varepsilon}{\mathfrak{d}} \left(C^{n+m} c_0 c'_0 + \sum_{l=1}^m \binom{n}{l} \binom{m}{l} \frac{l!}{\mu^{l-1}} \frac{\mu^{l-1} C^l}{n^l} c_l \right) \right).
\end{aligned}$$

Then,

$$\sum_{l=1}^m \binom{n}{l} \binom{m}{l} \frac{l!}{\mu^{l-1}} \frac{\mu^{l-1} C^l}{n^l} c_l \leq \sum_{l=1}^m \binom{m}{l} \frac{n!}{(n-l)! n^l} C^l c_l \ \leq (1 + C)^m \max_{1 \leq l \leq m} c_l$$

which concludes the proof.

- The inequalities (4.8) can be obtained in the same way. □

Note also the following bound in L^p norms of the fluctuation field.

Proposition 4.2. *For any $p \in [2, \infty)$, there exists a constant $C_p > 0$ such that*

$$\big(\mathbb{E}_\varepsilon \left[\zeta^0_\varepsilon(g)^p \right] \big)^{1/p} \leq C_p \|g\|_{L^p(M(v)dz)}. \tag{4.15}$$

The proof can be found in Appendix A of [BGSRS23b].

From these estimations, one can deduce the following corollary:

Corollary 4.3. *Let h_n be a test function satisfying the conditions of Proposition 4.1. Then there exists a constant $C > 0$ such that*

$$\left| \mathbb{E}_\varepsilon \left[\frac{1}{\sqrt{\mu}} \sum_{\underline{i}_n} h_n(\mathbf{Z}_{\underline{i}_n}(t_s)) \zeta^0_\varepsilon(g) \mathbb{1}_{\Upsilon_\varepsilon} \right] \right| \leq C^n \mu^{n-1} \mathbb{E}_\varepsilon \big[\zeta^0_\varepsilon(g)^2 \big]^{1/2} \left(c_0 + \left(\sup_{1 \leq l \leq n} c_l \right)^{1/2} \right). \tag{4.16}$$

Proof. We can decompose the left hand side as

$$\begin{aligned}\mathbb{E}_\varepsilon\left[\frac{1}{\sqrt{\mu}}\sum_{\underline{i}_n} h_n(\mathbf{Z}_{\underline{i}_n}(t_s))\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right] &= \mu^{n-1}\mathbb{E}_\varepsilon\left[\mu^{\frac12-n}\sum_{\underline{i}_n} h_n(\mathbf{Z}_{\underline{i}_n}(t_s))\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]\\ &= \mu^{n-1}\left(\mathbb{E}_\varepsilon\left[\mu^{\frac12}\widehat{h_n}(\mathbf{Z}_{\mathscr{N}}(t_s))\,\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right] + \mathbb{E}_\varepsilon\left[h_n\right]\mathbb{E}_\varepsilon\left[\mu^{\frac12}\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]\right)\\ &= \mu^{n-1}\left(\mathbb{E}_\varepsilon\left[\mu^{\frac12}\widehat{h_n}(\mathbf{Z}_{\mathscr{N}}(t_s))\,\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right] + \mathbb{E}_\varepsilon\left[h_n\right]\mathbb{E}_\varepsilon\left[\zeta^0_\varepsilon(g)\mu^{\frac12}\left(-\mathbb{1}_{\Upsilon^c_\varepsilon}\right)\right]\right).\end{aligned}$$

Using that $\mathbb{E}_\varepsilon[\zeta^0_\varepsilon(g)] = 0$. Applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}&\left|\mathbb{E}_\varepsilon\left[\mu^{-\frac12}\sum_{\underline{i}_n} h_n(\mathbf{Z}_{\underline{i}_n}(t_s))\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]\right|\\ &\qquad\le \mu^{n-1}\left(\mathbb{E}_\varepsilon\left[\mu\left[\widehat{h_n}\right]^2\right]^{\frac12}\mathbb{E}_\varepsilon\left[\zeta^0_\varepsilon(g)^2\right]^{\frac12} + \mathbb{E}_\varepsilon\left[h_n\right]\mathbb{E}_\varepsilon\left[\zeta^0_\varepsilon(g)^2\right]^{\frac12}\left(\mu\mathbb{P}_\varepsilon\left[\Upsilon^c_\varepsilon\right]\right)^{\frac12}\right).\end{aligned}$$

We apply now Proposition 4.1. The bound on $\mathbb{P}_\varepsilon\left[\Upsilon^c_\varepsilon\right]$ given in Section 3.2 and the bound on the L^p norm of $\zeta^0_\varepsilon(g)$ (4.15) lead to the stated corollary. □

5. Clustering estimations without recollision

The objective of this section is to bound $G^{\mathrm{clust}}_\varepsilon(t)$ and $G^{\exp}_\varepsilon(t)$, defined by

$$G^{\mathrm{clust}}_\varepsilon(t) := \mathbb{E}_\varepsilon\left[\zeta^t_\varepsilon(h)\zeta^0(g)\mathbb{1}_{\Upsilon^c_\varepsilon}\right] - \sum_{\substack{(n_j)_{j\le K}\\ 0\le n_j-n_{j-1}\le 2^j}} \mathbb{E}_\varepsilon\left[\frac{1}{\sqrt{\mu}}\sum_{\underline{i}_{n_K}}\Psi^{0,t}_{\underline{n}_K}[h]\left(\mathbf{Z}_{\underline{i}_{n_K}}(0)\right)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon^c_\varepsilon}\right],$$

$$G^{\exp}_\varepsilon(t) := \sum_{1\le k\le K}\ \sum_{\substack{(n_j)_{j\le k-1}\\ 0\le n_j-n_{j-1}\le 2^j}}\ \sum_{n_k\ge 2^k+n_{k-1}} \mathbb{E}_\varepsilon\left[\frac{1}{\sqrt{\mu}}\sum_{\underline{i}_{n_k}}\Psi^{0,k\theta}_{\underline{n}_k}[h]\left(\mathbf{Z}_{\underline{i}_{n_k}}(t-k\theta)\right)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right].$$

Proposition 5.1. *For $\varepsilon > 0$ small enough,*

$$|G^{exp}_\varepsilon(t) + G^{rec}_\varepsilon(t)| \le C\|g\|_0\|h\|_0\left(\varepsilon^{1/3}(Ct)^{2^{t/\theta}} + \tfrac{\theta t^2}{\mathfrak{d}^3}\right) \tag{5.1}$$

To obtain the stated result, we need first the following bounds on the pseudotrajectory developments without recollisions of type $\Psi^{0,k\theta}_{\underline{n}_k}[h]$:

Proposition 5.2. *Fix $k\in\mathbb{N}$, $\underline{n} := (n_1,\cdots,n_k)\in\mathbb{N}^k$ with $n_1\le n_2\le\cdots\le n_k$. Then, fixing $x_0 = 0$,*

$$\int \sup_{y\in\mathbb{T}}\left|\Psi^{0,k\theta}_{\underline{n}_k}[h](\mathrm{tr}_y\, Z_{n_k})\right|\frac{e^{-\mathscr{H}_{n_k}(Z_{n_k})}}{(2\pi)^{\frac{n_k d}{2}}}\,\mathrm{d}V_{n_k}\,\mathrm{d}X_{2,n_k} \le \frac{\|h\|_0}{(\mu\mathfrak{d})^{n_k-1}}C^{n_k}\theta^{n_k-n_{k-1}}(k\theta)^{n_{k-1}-1}, \tag{5.2}$$

and, for $m\in[1,n_k]$,

$$\begin{aligned}&\int \sup_{y\in\mathbb{T}}\left|\Psi^{0,k\theta}_{\underline{n}_k}[h]\circledast_m\Psi^{0,k\theta}_{\underline{n}_k}[h](\mathrm{tr}_y\, Z_{2n_k-m})\right|\frac{e^{-\mathscr{H}_{2n_k-m}(Z_{2n_k-m})}}{(2\pi)^{\frac{(2n_k-m)d}{2}}}\,\mathrm{d}V_{2n_K-m}\,\mathrm{d}X_{2,2n_K-m}\\ &\qquad\le \frac{\mu^{m-1}}{n_k^m}\left(\frac{\|h\|_0}{(\mu\mathfrak{d})^{n_k-1}}C^{n_k}\right)^2\theta^{n_k-n_{k-1}}(k\theta)^{n_{k-1}+n_k-1}.\end{aligned} \tag{5.3}$$

Using Corollary 4.3 and the previous estimations,

$$\begin{aligned}\left|\mathbb{E}_\varepsilon\left[\mu^{-\frac12}\sum_{\underline{i}_{n_k}}\Psi^{0,k\theta}_{\underline{n}_k}[h]\left(\mathbf{Z}_{\underline{i}_{n_k}}(t-k\theta)\right)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]\right| &\le \|g\|_0\|h\|_0C^{n_k}\left((\tfrac{\theta}{\mathfrak{d}})^{n_k-n_{k-1}}(\tfrac{k\theta}{\mathfrak{d}})^{n_{k-1}-1}\right.\\ &\qquad\left.+\,(\tfrac{\theta}{\mathfrak{d}})^{\frac{n_k-n_{k-1}}{2}}(\tfrac{k\theta}{\mathfrak{d}})^{\frac{n_k+n_{k-1}-1}{2}}\right)\\ &\le \|g\|_0\|h\|_0C^{n_k}(\tfrac{\theta}{\mathfrak{d}})^{\frac{n_k-n_{k-1}}{2}}(\tfrac{t}{\mathfrak{d}})^{n_k},\end{aligned}$$

and in the same way,

$$\mathbb{E}_\varepsilon\left[\mu^{-\frac12}\sum_{\underline{i}_{n_K}}\Psi^{0,k\theta}_{\underline{n}_k}[h]\left(\mathbf{Z}_{\underline{i}_{n_K}}(0)\right)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon^c_\varepsilon}\right]=O\left(\varepsilon^{\frac12}\|g\|_0\|h\|_0C^{n_k}(t/\mathfrak{d})^{n_k}\right).$$

Summing over all possible $(n_1,\cdots,n_k)$, we obtain

$$\begin{aligned}|G^{\exp}_\varepsilon(t)|&\le\sum_{k=1}^K\sum_{\substack{n_1\le\cdots\le n_{k-1}\\ n_j-n_{j-1}\le 2^j}}\sum_{n_k>2^k+n_{k-1}}\|g\|_0\|h\|_0C^{n_k}(\theta/\mathfrak{d})^{\frac{n_k-n_{k-1}}{2}}(t/\mathfrak{d})^{n_k}\\&\le\|g\|_0\|h\|_0\sum_{k=1}^K\sum_{\substack{n_1\le\cdots\le n_{k-1}\\ n_j-n_{j-1}\le 2^j}}\sum_{n_k>2^k+n_{k-1}}\left(C\frac{\theta t^2}{\mathfrak{d}^3}\right)^{\frac{n_k-n_{k-1}}{2}}\\&\le C\|g\|_0\|h\|_0\sum_{k=1}^K2^{k^2}\left(C\frac{\theta t^2}{\mathfrak{d}^3}\right)^{2^{k-1}}\le C\|g\|_0\|h\|_0\frac{\theta t^2}{\mathfrak{d}^3}.\end{aligned}$$

as the series converges for θ small enough. In the same way

$$\begin{aligned}\left|G^{\mathrm{clust}}_\varepsilon(t)\right|\le&\mathbb{P}_\varepsilon(\Upsilon^c_\varepsilon)^{\frac14}\mathbb{E}_\varepsilon[\zeta^0_\varepsilon(g)^4]^{\frac14}\mathbb{E}_\varepsilon[\zeta^0_\varepsilon(h)^2]^{\frac12}\\&+\sum_{\substack{n_1\le\cdots\le n_K\\ n_j-n_{j-1}\le 2^j}}\mathbb{P}_\varepsilon(\Upsilon^c_\varepsilon)^{\frac14}\|g\|_0\|h\|_0C^{n_k}t^{n_k}+\|g\|_0\|h\|_0(\tfrac{C\theta}{\mathfrak{d}})^{n_K}(\tfrac{\varepsilon}{\mathfrak{d}})^{\frac12}\\\le&C\|g\|_0\|h\|_0\varepsilon^{\frac13}2^{K^2}(Ct)^{2^K}.\end{aligned}$$

This concludes the proof of (5.1).

Proof of (5.2). We recall that for $t=k\theta$ and that

$$\Psi^{0,k\theta}_{\underline{n}_k}[h]:=\frac{1}{(n_k-1)!}\sum_{(s_l)_{l\le n_k-1}}\prod_{l=1}^{n_k-1}s_lh(\mathsf{z}_1(t,\cdot,\{1\},(s_l)_l))\mathbb{1}_{\mathscr{R}^{0,t}_{\{q\},(s_l)_l}}\prod_{i=0}^{k-1}\mathbb{1}_{\mathfrak{n}(i\theta)=n_{n_{k-i}}}.$$

This gives directly the following bound on $\bar\Psi^{0,t}_{\underline{n}_k}[h]$

$$\left|\bar\Psi^{0,t}_{\underline{n}_k}[h]\right|\le\frac{\|h\|_0}{(n_k-1)!}\sum_{(s_l)_{l\le n_k-1}}\sum_{(s_l)_{l\le n_k-1}}\mathbb{1}_{\mathscr{R}^{0,t}_{\{1\},(s_l)_l}}\mathbb{1}_{\mathfrak{n}(\theta)=n_{k-1}}.\tag{5.4}$$

As the right-hand side of (5.4) is invariant under translations, it is sufficient to fix $x_1=0$ and integrate with respect to (X_{2,n_k},V_{n_k}).

We define the *clustering tree* $T^>$ as the sequence $(q_i,\bar q_i)_{1\le i\le n_k-1}$ where the i-th collision involves particles q_i and $\bar q_i$ (and $q_i<\bar q_i$).

Since in the present section, pseudotrajectories have no recollision, the collision graph is a tree (a simply connected graph). Hence, we can construct the clustering by forgetting the scattering times associated with each edge but keeping the order of the collisions. It can be used to parametrize a partition of $\mathscr{R}^{0,t}_{\{q\},(s_l)_l}$.

Let us fix a clustering tree. We perform the following change of variables

$$X_{2,n_k}\mapsto(\hat x_1,\cdots,\hat x_{n_k-1}),\ \forall i\in[1,n_k-1],\ \hat x_i:=x_{q_i}-x_{\bar q_i}$$

Fix then τ_{i+1} the time of the $(i+1)$-th collision, as well as the relative positions $\hat x_1,\cdots,\hat x_{i-1}$. We denote $T_i=\theta$ if $i\le n_k-n_{k-1}$, t else (at least n_k-n_{k-1} clustering collisions happen before time θ). The i-th collision set is defined by

$$B_{T^>,i}:=\left\{\hat x_i\middle|\exists\tau\in(0,T_i\wedge\tau_{i+1}),\ |\mathsf{x}_{q_i}(\tau)-\mathsf{x}_{\bar q_i}(\tau)|\le\varepsilon\right\}.$$

Because particles $\mathsf{x}_{q_i}(\tau)$ and $\mathsf{x}_{\bar q_i}(\tau)$ are independent until their first meeting, we can perform the change of variable $\hat x_i\mapsto(\tau_i,\eta_i)$ where τ_i is the first meeting time and

$$\eta_i:=\frac{\mathsf{x}_{q_i}(\tau_i)-\mathsf{x}_{\bar q_i}(\tau_i)}{|\mathsf{x}_{q_i}(\tau_i)-\mathsf{x}_{\bar q_i}(\tau_i)|}.$$

It sends the Lebesgue measure $\mathrm{d}\hat{x}_i$ to the measure $\varepsilon^{d-1}((\mathsf{v}_{q_i}(\tau_i)-\mathsf{v}_{\bar{q}_i}(\tau_i))\cdot\eta_i)_+\,\mathrm{d}\eta_i\,\mathrm{d}\tau_i$ and

$$\int \mathbb{1}_{B_{T^>,i}}\,\mathrm{d}\hat{x}_i \le C\varepsilon^{d-1}\int_0^{T_i\wedge\tau_{i+1}} |\mathsf{v}_{q_i}(\tau_i)-\mathsf{v}_{\bar{q}_i}(\tau_i)|\,\mathrm{d}\tau_i.$$

We want now to sum on every possible edge $(q_i,\bar{q}_i)$. Hence, we need to control

$$\sum_{(q_i,\bar{q}_i)} |\mathsf{v}_{q_i}(\tau_i)-\mathsf{v}_{\bar{q}_i}(\tau_i)| \le 2n_k\sum_k |\mathsf{v}_k(\tau_i)| \le 2n_k\left(n_k\sum_k|\mathsf{v}_k(\tau_i)|^2\right)^{1/2} \le n_k\left(n_k+|\mathsf{V}_{n_k}(\tau_i)|^2\right)$$

Lemma 5.3. *Consider a time $\tau\in[0,t]$, collision parameters $(\omega_1,\omega_2,(s_i)_i)$ and an initial position $Z_n\in\mathbb{D}^n$. Then*

$$\tfrac{1}{2}|\mathsf{V}(\tau,Z_n,\omega_1,\omega_2,(s_i)_i)|^2 \le \mathscr{H}_n(Z_n),$$

as there is no overlap between particles.

Proof. We begin by defining the notion of *clusters on a time segment*:

Definition 5.1. *Consider two times $0\le\tau_a<\tau_b\le t$. We denote $\mathscr{G}$ the collision graph of the pseudotrajectory $\mathsf{Z}_{n_k}(\cdot,(\omega,(s_i)_i),Z_{n_k})$ on the time interval $[\tau_a,\tau_b]$ and G a graph with edges*

$$\Big\{(q,\bar{q})\in[n]^2,\ \exists\tau'\in[\tau_a,\tau_b],\ (q,\bar{q})_{1,\tau}\in\mathscr{G}\}.$$

We take only into account the collisions with interaction. We define $\underline{\kappa}:=(\kappa_1,\cdots,\kappa_k)$ the clusters on the segment $[\tau_1,\tau_2]$ *the connected components of G (defined in the following of Definition 3.8).*

Note that if τ_a lies between the beginning of the collision implying s_j and the beginning of the collision implying s_{j+1}, then $\underline{\kappa}$ only depends on the $(s_i)_{i\le j}$.

We distinguished the cases $\tau>\delta$ and $\tau\le\delta$.

- First, if $\tau\le\delta$. We consider $(\kappa_1,\cdots,\kappa_k)$ the cluster on the segment $[0,\delta]$ constructed in the following Definition 3.8. The pseudotrajectory is the Hamiltonian trajectory associated with the energy

$$\mathscr{H}_{\underline{\kappa}}(Z_n):=\sum_{i=1}^k\left(\sum_{q\in\kappa_i}\frac{|v_q|^2}{2}+\sum_{\substack{q,\bar{q}\in\kappa_i\\ q\neq\bar{q}}}\frac{\alpha}{2}\mathscr{V}\left(\frac{x_q-x_{\bar{q}}}{\varepsilon}\right)\right).$$

  Hence

$$\tfrac{1}{2}|\mathsf{V}(\tau,Z_n,\omega_1,\omega_2,(s_i)_i)|^2 \le \mathscr{H}^{\underline{\kappa}}(\mathsf{Z}_n(\tau)\le\mathscr{H}^{\underline{\kappa}}(Z_n)\le\mathscr{H}_n(Z_n).$$

- If $\tau>\delta$, consider $\underline{\kappa}$ and $\underline{\kappa}'$ the clusters on $[\delta,\tau]$ and on $[0,\delta]$. After time δ, the particles outside ω_2 stop interacting, and before time δ, the couple of particles in ω_2 cannot overlap. Hence, $\underline{\kappa}'$ is a finer partition of $[n]$ than $\underline{\kappa}$ and $\mathscr{H}_{\underline{\kappa}}\le\mathscr{H}_{\underline{\kappa}'}$. Thus

$$\frac{1}{2}|\mathsf{V}_{n_k}(\tau)|^2 \le \mathscr{H}_{\underline{\kappa}}(\mathsf{Z}_{n_k}(\tau))=\mathscr{H}_{\underline{\kappa}}(\mathsf{Z}_{n_k}(\delta))\le\mathscr{H}_{\underline{\kappa}'}(\mathsf{Z}_{n_k}(\delta))=\mathscr{H}_{\underline{\kappa}'}(\mathsf{Z}_{n_k}(0))\le\mathscr{H}_n(Z_n). \tag{5.5}$$

□

As we suppose that there is no overlap at time 0, we have $\mathscr{H}_n(Z_n)=\frac{|V_n|^2}{2}$. Hence, using the Boltzmann-Grad scaling $\mu\varepsilon^{d-1}\mathfrak{d}=1$,

$$\begin{aligned}
&\sum_{(q_i,\bar{q}_i)_i}\int\mathrm{d}\hat{x}_1\mathbb{1}_{B_{T^>,1}}\cdots\int\mathrm{d}\hat{x}_{n_k-1}\mathbb{1}_{B_{T^>,n_k-1}}\mathbb{1}_{\mathfrak{n}(\theta)=n_{k-1}}e^{-\mathscr{H}_{n_k}(Z_{n_k})}\\
&\quad\le\left(\frac{Cn_k}{\mu\mathfrak{d}}\right)^{n_k-1}(n_k+\mathscr{H}_{n_k}(Z_{n_k}))^{n_k-1}e^{-\mathscr{H}_{n_k}(Z_{n_k})}\int_0^{T_{n_k-1}}\mathrm{d}\tau_{n_k-1}\cdots\int_0^{T_1\wedge\tau_2}\mathrm{d}\tau_1\\
&\quad\le\left(\frac{Cn_k}{\mu\mathfrak{d}}\right)^{n_k-1}n_k^{n_k-1}e^{-\frac{\mathscr{H}_{n_k}(Z_{n_k})}{2}}\frac{t^{n_{k-1}-1}}{(n_{k-1}-1)!}\frac{\theta^{n_k-n_{k-1}}}{(n_k-n_{k-1})!}\\
&\quad\le\left(\frac{\tilde{C}}{\mu\mathfrak{d}}\right)^{n_k-1}n_k^{n_k-1}e^{-\frac{|V_{n_k}|^2}{4}}t^{n_{k-1}-1}\theta^{n_k-n_{k-1}},
\end{aligned}$$

We used these two classical inequalities

$$\frac{(a+b)^{a+b}}{a!b!} \leq e^{a+b}\frac{(a+b)!}{a!b!} \leq (2e)^{a+b}$$

$$\forall A,B>0,\ ,x\in\mathbb{R}^+,\quad (A+x)^B e^{-\frac{x}{2}} = B^B\left(\frac{A+x}{B}e^{-\frac{A+x}{2B}}\right)^B e^{\frac{A}{4}} \leq \left(\tfrac{4B}{e}\right)^B e^{\frac{A}{4}}.$$

Finally, we sum on V_{n_k}, on the 2^{n_k-1} possible $(s_i)_i$ and on the $q\in[1,n_k]$, and we divide by the remaining $(n_k!)$. This gives the expected estimation. □

Proof of (5.3). We begin as in the previous paragraph. Using that

$$\left|\left\{(\omega,\omega',q,q')\right|\ \omega\cup\omega'=[2n_k-m],\ |\omega|=|\omega'|=n_k,\ q\in\omega,\ q'\in\omega'\right\}| = \frac{n_k^2(2n_k-m)!}{(n_k-m)!^2m!},$$

and that the right hand side of (5.4) is symmetric, one has

$$\begin{aligned}(5.6)\quad &\left|\Psi^{0,t}_{\underline{n}_k}[h]\circledast_m\bar\Psi^{0,t}_{\underline{n}_k}[h](Z_{2n_k-m})\right|\\ &\leq \frac{\|h\|_0^2}{(n_k!)^2}\frac{(n_k-m)!^2m!}{n_k^2(2n_k-m)!}\sum_{(\omega,\omega',q,q')}\ \sum_{\substack{(s_l)_{l\leq n_k-1}\\(s'_l)_{l\leq n_k-1}}}\mathbb{1}_{\mathscr{R}^{0,t}_{\{q\},(s_l)_l}}(Z_\omega)\mathbb{1}_{\mathfrak{n}(\theta)=n_{k-1}}(Z_\omega)\mathbb{1}_{\mathscr{R}^{0,t}_{\{q'\},(s'_l)_l}}(Z_{\omega'}).\end{aligned}$$

where $\mathfrak{n}(\theta)$ is the number of particles at time θ in the pseudotrajectory $\mathsf{Z}(t,\cdot,\{1\},(s_l)_l)$. The right-hand side is invariant under translation. Hence, without loss of generality, we can suppose that $1\notin\omega'$ and fix $x_1=0$.

We have to consider two pseudotrajectories

$$\mathsf{Z}(\tau):=\mathsf{Z}(\tau,Z_\omega,\{q\},(s_l)_l) \text{ and } \mathsf{Z}'(\tau):=\mathsf{Z}(\tau,Z_{\omega'},\{q'\},(s'_l)_l).$$

We want to estimate

$$\int \mathbb{1}_{\mathscr{R}^{0,t}_{\{q'\},(s'_l)_l}}(Z_{\omega'})e^{-\frac12\mathscr{H}_{2n_k-m}(Z_{2n_k-m})}\,\mathrm{d}Z_{\omega'\setminus\omega}.$$

Fix Z_ω and denote T_a the clustering tree of the pseudotrajectory $\mathsf{Z}(t)$, constructed as in the proof of (5.2). Next, we construct the clustering tree associated with the second pseudotrajectory: let $(q_i,\bar q_i)_{i\leq\ell}$ be the edges of the collision graph of $\mathsf{Z}'(\tau)$, taking temporal order. Set $\bar T_0=\emptyset$. Suppose that $\bar T_i$ is constructed. Then $\bar T_{i+1}:=\bar T_i\cup\{(q_i,\bar q_i)\}$ if the graph $T_a\cup\bar T_i\cup\{(q_i,\bar q_i)\}$ has no cycle. Else $\bar T_{i+1}:=\bar T_i$. At the ℓ-step we have constructed an ordered graph $T_b:=\bar T_\ell$ with n_k-m edges.

The T_b define a partition of $\{Z_{\omega'\setminus\omega}\in\mathbb{D}^{n_k-m}|Z_{\omega'}\in\mathscr{R}^{0,t}_{\{q'\},(s'_l)_l}\}$ where the coordinates Z_ω are fixed.

The rest of the proof is almost identical to the proof of (5.2). Fix the clustering tree $T_b=(q_i,\bar q_i)_{n_k\leq i\leq 2n_k-m}$, and perform the following change of variables

$$X_{\omega'\setminus\omega}\mapsto(\hat x_{n_k},\cdots,\hat x_{2n_k-m-1}),\ \forall i\in[n_k,2n_k-m-1],\ \hat x_i:=x_{q_i}-x_{\bar q_i}.$$

Fix τ_{i+1}, the time of the $(i+1)$-th collision, and relative positions $\hat x_{n_k},\cdots,\hat x_{i-1}$. We define the i-th collision set as

$$B_{T^>,i}:=\left\{\hat x_i\middle|\exists\tau\in(0,t\wedge\tau_{i+1}),\ |\mathsf{x}'_{q_i}(\tau)-\mathsf{x}'_{\bar q_i}(\tau)|\leq\varepsilon\right\}.$$

As in the preceding lemma, we can perform the change of variable $\hat x_i\mapsto(\tau_i,\eta_i)$ where τ_i is the first meeting time and

$$\eta_i:=\frac{\mathsf{x}'_{q_i}(\tau_i)-\mathsf{x}'_{\bar q_i}(\tau_i)}{\varepsilon}.$$

We have

$$\sum_{(q_i,\bar q_i)}\int\mathbb{1}_{B_{T^>,i}}\,\mathrm{d}\hat x_i\leq C\varepsilon^{d-1}\int_0^{\tau_{i+1}}\sum_{(q_i,\bar q_i)\in\omega'^2}\left|\mathsf{v}'_{q_i}(\tau_i)-\mathsf{v}'_{\bar q_i}(\tau_i)\right|\,\mathrm{d}\tau_i.$$

Using the same method as in the proof of (5.5), we have

$$\sum_{(q_i,\bar q_i)\in\omega'^2}\left|\mathsf{v}'_{q_i}(\tau_i)-\mathsf{v}'_{\bar q_i}(\tau_i)\right|\leq n_k+|\mathsf{V}'_{\omega'}(\tau_i)|^2\leq 2n_k+2\mathscr{H}_{n_k}(Z_{\omega'}).$$

Using that

$$\mathscr{H}_{2n_k-m}(Z_{2n_k-m})\leq\tfrac12\left(\mathscr{H}_{n_k}(Z_\omega)+\mathscr{H}_{n_k}(Z_{\omega'})\right),$$

we can apply the same computation as above,

$$\int \mathbb{1}_{\mathscr{R}^{0,t}_{\{q'\},(s'_l)_l}}(Z_{\omega'})e^{-\mathscr{H}_{2n_k-m}}\,\mathrm{d}Z_{\omega'\setminus\omega} \le \sum_{T_b}\int e^{-\frac{\mathscr{H}_{n_k}(Z_\omega)}{2}}\prod_{i=1}^{n_k-m}\mathbb{1}_{B_{T^>,i}}\,\mathrm{d}\hat{x}_i\; e^{-\mathscr{H}_{n_k}(Z_{\omega'})}dV_{\omega'}$$
$$\le C\left(\frac{C}{\mu\mathfrak{d}}\right)^{n_k-m} t^{n_k-m}(2n_k-m)^{n_k-m}.$$

We can estimate

$$\int \mathbb{1}_{\mathscr{R}^{0,t}_{\{q\},(s_l)_l}}(Z_\omega)e^{-\frac{1}{2}\mathscr{H}_{2n_k-m}}\,\mathrm{d}X_{\omega\setminus\{1\}}\,\mathrm{d}V_\omega$$

as in the proof of (5.2). We get the expected result by summing on all the possible parameters $(s_i)_i$, $(s'_i)_i$, q, q', ω and ω'. □

6. Treatment of the main part

The aim of this section is the proof of

$$G^{\mathrm{main}}_\varepsilon(t) = \int_{\mathbb{D}} h(z)\mathbf{g}_\alpha(t,z)M(z)dz + O\left(\left(C\tfrac{\theta t}{\mathfrak{d}^2}+\varepsilon^{\mathfrak{a}}(\tfrac{Ct}{\mathfrak{d}})^{2^{K+1}}\right)\|h\|_1\|g\|_1\right),$$

where $\mathbf{g}_\alpha(t,z)$ is the solution of the linearized Boltzmann equation (2.16) and $\mathfrak{a}\in(0,1)$ is some fix constant depending only on the dimensio n.

6.1. Duality formula.

We recall that

$$G^{\mathrm{main}}_\varepsilon(t) = \sum_{\substack{(n_j)_{j\le K}\\ 0\le n_j-n_{j-1}\le 2^j}} \mathbb{E}_\varepsilon\left[\frac{1}{\sqrt{\mu}}\sum_{\underline{i}_{n_K}}\Psi^{0,t}_{\underline{n}_K}[h]\left(\mathbf{Z}_{\underline{i}_{n_K}}(0)\right)\zeta^0_\varepsilon(g)\right]$$
$$= \sum_{\substack{n_1\le\cdots\le n_K\\ n_j-n_{j-1}\le 2^j}} \mathbb{E}_\varepsilon\left[\mu^{n_K}\,\hat{\Psi}^{0,t}_{\underline{n}_K}[h]\,\hat{g}\right]$$

where $\Psi^{0,t}_{\underline{n}_K}[h]$ is the development of $h(z_i(t))$ along pseudotrajectories with n_k remaining particles at time $t-k\theta$, and neither recollision nor overlap nor multiple encounters .

We denote

$$g_{n_K}(Z_{n_K}) := \sum_{i=1}^{n_K} g(z_i) \tag{6.1}$$

Then, using the equality (4.7) and L^1 estimations on $\Psi^{0,t}_{\underline{n}_K}[h]$ of Section 5, we have for h and g in L^∞

$$G^{\mathrm{main}}_\varepsilon(t)$$
$$= \sum_{\substack{n_1\le\cdots\le n_K\\ n_j-n_{j-1}\le 2^j}}\int \mu^{n_K-1}\Psi^{0,t}_{\underline{n}_K}[h]\,(Z_{n_K})\,g_{n_K}(Z_{n_K})\,\frac{e^{-\mathscr{H}_{n_K}(Z_{n_K})}\,\mathrm{d}Z_{n_K}}{(2\pi)^{\frac{n_K d}{2}}} + O\left(\tfrac{\varepsilon}{\mathfrak{d}}\sum_{\underline{n}}(\tfrac{Ct}{\mathfrak{d}})^{n_k}\|h\|_0\|g\|_0\right)$$
$$= \sum_{\underline{n}_K}\int \mu^{n_K-1}\Psi^{0,t}_{\underline{n}_K}[h]\,(Z_{n_K})\,g_{n_K}(Z_{n_K})\,\frac{e^{-\mathscr{H}_{n_K}(Z_{n_K})}\,\mathrm{d}Z_{n_K}}{(2\pi)^{\frac{n_K d}{2}}} + O\left(\tfrac{\varepsilon}{\mathfrak{d}}\left(K2^{K^2}(\tfrac{Ct}{\mathfrak{d}})^{2^{K+1}}\|h\|_0\|g\|_0\right)\right).$$

We want to compute the asymptotics of each term in the sum. As we suppose that there is no overlap

$$\int \mu_\varepsilon^{n_K-1}\Psi^{0,t}_{\underline{n}_K}[h]\,(Z_{n_K})\,g_{n_K}(Z_{n_K})\frac{e^{-\mathscr{H}_{n_K}(Z_{n_K})}\,\mathrm{d}Z_{n_K}}{(2\pi)^{\frac{n_K d}{2}}}$$
$$= \frac{\mu^{n_K-1}}{(n_K-1)!}\sum_{(s_k)_k}\prod_{k=1}^{n_l-1}s_k\int_{\mathscr{R}^{0,t}_{\{1\},(s_k)_k}} h(\mathsf{z}^\varepsilon_1(t,Z_{n_K},\{1\},(s_k)_k))g_{n_K}(Z_{n_K})\prod_{i=1}^{K}\mathbb{1}_{\mathfrak{n}(t-i\theta)=n_i}M^{\otimes n_K}\,\mathrm{d}Z_{n_K}.$$

where $\mathscr{R}^{0,t}_{\{1\},(s_k)_k}$ is the set of initial parameters such that the pseudotrajectory has no recollision and $\mathfrak{n}(\tau)$ is the number of remaining particles at time τ (see definition 3.12). We had an exponent ε on z^ε_1 to mark the ε-dependence of the pseudotrajectory.

We want to construct the limiting process of the pseudotrajectory $\mathsf{Z}^\varepsilon_n(\tau)$.

We denote T the *clustering tree* as the sequence $(q_i, \bar{q}_i, \bar{s}_i)_{i\le n_K-1}$ such that the i-th collision happens between particles q_i and $\bar{q}_i$ (with $q_i < \bar{q}_i$) and $\bar{s}_i$ is equal to 1 (respectively -1) if the particles interact (respectively do not interact). Fixing the initial velocities V_{n_K}, we perform the change of variable

$$X_{n_K} \mapsto (x_1, (\nu_i, \tau_i)_{i\le n_K-1}), .$$

where τ_i is the first time when particles q_i and $\bar{q}_i$ meet, an

$$\nu_i = \frac{\mathsf{x}^\varepsilon_{q_i}(\tau_i) - \mathsf{x}^\varepsilon_{\bar{q}_i}(\tau_i)}{\varepsilon}.$$

The Jacobian of this application is

$$\mathrm{d}X_{n_K} \to \prod_{i=1}^{n_k-1} \varepsilon^{d-1} \left(\left(\mathsf{v}^\varepsilon_{q_i}(\tau_i) - \mathsf{v}^\varepsilon_{\bar{q}_i}(\tau_i) \right) \cdot \nu_i \right)_+ \mathrm{d}\nu_i \, \mathrm{d}\tau_i =: \frac{\Lambda_T(V_{n_K}, \nu_{[n_k-1]})}{(\mu\mathfrak{d})^{n_k-1}} \, \mathrm{d}\nu_{[n_k-1]} \, \mathrm{d}\tau_{[n_K-1]} \; \mathrm{d}x_1$$

where we have denoted

$$\nu_{[n_k-1]} = (\nu_1, \cdots, \nu_{n_k-1}), \quad \tau_{[n_k-1]} = (\tau_1, \cdots, \tau_{n_k-1}).$$

The kernel $\Lambda(V_{n_K}, \nu_{[n_k-1]})$ only depends on the successive velocities $(\mathsf{v}^\varepsilon_{q_i}(\tau_i), \mathsf{v}^\varepsilon_{\bar{q}_i}(\tau_i))$ which can be deduced from the collision graph, forgetting the exact values of the $\mathrm{d}\tau_{[n_k-1]}$ (since we have forbidden the pathological pseudotrajectories).

We defined the signature of the collision tree $\sigma(T) := \bar{s}_1 \bar{s}_2 \cdots \bar{s}_{n_K}$, the set of collision times

$$\mathfrak{T}_{\underline{n}_K} := \{(\tau_i)_{i\le n_K-1}, \ \tau_i \le \tau_{i+1}, \ \forall k \le K, \ j \in [n_K - n_{K-k}, n_K - n_{K-k-1}], \ k\theta \le \tau_j \le (k+1)\theta\}$$

and for a given family $\tau_{[n_K-1]}$, we define $\mathfrak{G}^\varepsilon_T(\tau_{[n_K-1]})$ the set of coordinates $(x_1, (\nu_i)_{i\le n_K-1}, V_{n_K})$ such that the pseudotrajectory has no recollision, and for all j, $\left(\mathsf{v}^\varepsilon_{q_i}(\tau_i) - \mathsf{v}^\varepsilon_{\bar{q}_i}(\tau_i)\right) \cdot \nu_i$ is positive.

The map

$$\bigsqcup_{(s_k)_k} \{s_k\} \times \left(\mathscr{R}^{0,t}_{\{1\},(s_k)_k} \cap \{\text{no overlap}\} \cap \bigcap_{j\le K-1} \{\mathfrak{n}(j\theta) = n_{K-j}\} \right) \to \bigsqcup_T \{T\} \times \mathfrak{T}_{\underline{n}_K} \times \mathfrak{G}^\varepsilon_T$$
$$(X_{n_K}, V_{n_K}) \mapsto (x_1, (\nu_i, \tau_i)_{i\le n_K-1}, V_{n_K})$$

is a diffeomosphism and

$$\begin{aligned} (6.2) \quad & \int \mu^{n_K-1} \Psi^{0,t}_{\underline{n}_K}[h]\,(Z_{n_K})\, g_{n_K}(Z_{n_K}) M^{\otimes n_K} \, \mathrm{d}Z_{n_K} \\ & = \frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!} \sum_T \sigma(T) \int_{\mathfrak{T}_{\underline{n}_K} \times \mathfrak{G}^\varepsilon_T} h(\mathsf{z}^\varepsilon_1(t,T) g_{n_K}(\mathsf{Z}^\varepsilon_{n_K}(0,T)) \\ & \qquad \times M^{\otimes n_K} \Lambda_T(V_{n_K}, \nu_{[n_k-1]}) \, \mathrm{d}\nu_{[n_k-1]} \, \mathrm{d}\tau_{[n_K-1]} \, \mathrm{d}x_1 \, \mathrm{d}V_{n_K}. \end{aligned}$$

Definition 6.1 (Pseudotrajectories for punctual particles)**.** *Fix a collision tree $T := (q_i, \bar{q}_i, \bar{s}_i)$ and collision parameters $(V_{n_k}, \tau_{[n_k-1]}, \nu_{[n_k-1]})$. We now define the pseudotrajectories for punctual particles. The velocities $\mathsf{V}^0_{n_K}(\tau, T)$ follow a jump process: at time 0, $\mathsf{V}^0_{n_K}(\tau = 0, T) = V_{n_K}$. At time τ_i, if $\bar{s}_i = 1$ the velocities of particles q_i, $\bar{q}_i$ jump to $v_{q_i}(\tau_i^+)$, $v_{\bar{q}_i}(\tau_i^+)$ given by $(v_{q_i}(\tau_i^+), v_{\bar{q}_i}(\tau_i^+), \tilde{\nu}_i) := \xi_\alpha(v_{q_i}(\tau_i^-), v_{\bar{q}_i}(\tau_i^-), \nu_i)$ (ξ_α the scattering map defined in Definition 2.1).*

We defined $\mathfrak{G}^0_T$ the set of the $(x_1, (\nu_i)_{i\le n_K-1}, V_{n_K})$ such that for all j, $\left(\mathsf{v}^0_{q_i}(\tau_i) - \mathsf{v}^0_{\bar{q}_i}(\tau_i)\right) \cdot \nu_i$ is positive. Note that $\mathfrak{G}^\varepsilon_T \subset \mathfrak{G}^0_T$.

We have formally the convergence

$$\begin{aligned} & \int \mu^{n_K-1} \Psi^{0,t}_{\underline{n}_K}[h]\,(Z_{n_K})\, g_{n_K}(Z_{n_K}) M^{\otimes n_K} \, dZ_{n_K} \\ & \qquad \xrightarrow[\varepsilon\to 0]{} \frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!} \sum_T \sigma(T) \int_{\mathfrak{T}_{\underline{n}_K} \times \mathfrak{G}^0_T} h(\mathsf{z}^0_1(t,T) g_{n_K}(\mathsf{Z}^0_{n_K}(0,T)) \\ & \qquad\qquad \times \Lambda_T(V_{n_K}, \nu_{[n_k-1]}) \, \mathrm{d}\nu_{[n_k-1]} \, \mathrm{d}\tau_{[n_K-1]} \, \mathrm{d}x_1 M^{\otimes n_K} \, \mathrm{d}V_{n_K}. \end{aligned}$$

In order to have explicit rates of convergence, we decompose the error into two parts:

$$\int \mu^{n_K-1}\,\Psi^{0,t}_{\underline{n}_K}[h]g^{\varepsilon}_{n_K}M^{\otimes n_K}\,\mathrm{d}Z_{n_K} \tag{6.3}$$
$$= (6.4) + (6.5) + \frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!}\sum_T \sigma(T)\int_{\mathfrak{T}_{\underline{n}_K}\times\mathfrak{G}^0_T} h(\mathsf{z}^0_1(t,T)g_{n_K}(\mathsf{Z}^0_{n_K}(0,T))$$
$$\times \Lambda_T(V_{n_K},\nu_{[n_k-1]})\,\mathrm{d}\nu_{[n_k-1]}\,\mathrm{d}\tau_{[n_K-1]}\,\mathrm{d}x_1 M^{\otimes n_K}\,\mathrm{d}V_{n_K},$$

where we define

$$= \frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!}\sum_T \sigma(T)\int_{\mathfrak{T}_{\underline{n}_K}\times\mathfrak{G}^{\varepsilon}_T}\Big(h(\mathsf{z}^{\varepsilon}_1(t,T)g_{n_K}(\mathsf{Z}^{\varepsilon}_{n_K}(0,T)) - h(\mathsf{z}^0_1(t,T)g_{n_K}(\mathsf{Z}^0_{n_K}(0,T))\Big) \tag{6.4}$$
$$\times \Lambda_T(V_{n_K},\nu_{[n_k-1]})\,\mathrm{d}\nu_{[n_k-1]}\,\mathrm{d}\tau_{[n_K-1]}\,\mathrm{d}x_1 M^{\otimes n_K}\,\mathrm{d}V_{n_K},$$

$$= -\frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!}\sum_T \sigma(T)\int_{\mathfrak{G}^0_T} h(\mathsf{z}^{\varepsilon}_1(t,T)g_{n_K}(\mathsf{Z}^{\varepsilon}_{n_K}(0,T))\left(1-\mathbb{1}_{\mathfrak{G}^{\varepsilon}_T}\right) \tag{6.5}$$
$$\times \Lambda_T(V_{n_K},\nu_{[n_k-1]})\,\mathrm{d}\nu_{[n_k-1]}\,\mathrm{d}\tau_{[n_K-1]}\,\mathrm{d}x_1 M^{\otimes n_K}\,\mathrm{d}V_{n_K},$$

The error parts are estimated using the following standard results:

Lemma 6.1. *Fix $\bar{n} := (n_1,\cdots,n_k)$ and denote for $p\in[1,2]$*

$$\Lambda^p_T(V_{n_K},\nu_{[n_k-1]}) := \prod_{i=1}^{n_k-1}\left|\mathsf{v}^0_{q_i}(\tau_i^-) - \mathsf{v}^0_{\bar{q}_i}(\tau_i^-)\right|^p$$

For any $\varepsilon>0$ sufficiently small, we have

$$\frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!}\sum_T \sigma(T)\int_{\mathfrak{T}_{\underline{n}_K}\times\mathfrak{G}^0_T}\Lambda^p_T(V_{n_K},\nu_{[n_K-1]})\,\mathrm{d}\tau_{[n_K-1]}\,\mathrm{d}\nu_{[n_K-1]}\,\mathrm{d}x_1 M^{\otimes n_K}\,\mathrm{d}V_{n_K} \tag{6.6}$$
$$\le C^{n_K}t^{n_{K-1}-1}\theta^{n_K-n_{K-1}}.$$

Proof. Fix first the collision tree $T := (q_i,\bar{q}_i,\bar{s}_i)_i$. We sum on each ν_i in the decreasing order:

$$\sum_{(q_i,\bar{q}_i,\bar{s}_i)}\int\left|\mathsf{v}^0_{q_i}(\tau_i^-) - \mathsf{v}^0_{\bar{q}_i}(\tau_i^-)\right|^p\,\mathrm{d}\nu_i \le C\sum_{(\bar{q}_i,\bar{s}_i)}|\mathsf{v}_{q_i}(\tau_i) - \mathsf{v}_{q_i}(\tau_i)|^p \le Cn_K^{2-\frac{p}{2}}\left|\mathsf{V}_{n_K}(\tau_i)\right|^{\frac{p}{2}} \tag{6.7}$$
$$\le Cn_K^{2-\frac{p}{2}}\left|V_{n_K}\right|^{\frac{p}{2}}$$

using the conservation of energy.

Hence,

$$\sum_T\int\Lambda^p_T(V_{n_K},\nu_{[n_k-1]})\,\mathrm{d}\nu_{[n_k-1]}M^{\otimes n_K}dV_{n_K} \le C^{n_K}n_K^{2-\frac{p}{2}}\int|V_{n_K}|^{\frac{n_Kp}{2}}\,M^{\otimes n_K}\,\mathrm{d}V_{n_K}$$
$$\le C_1^{n_K}n_K^{2n_K}\int e^{-\frac{1}{4}|V_{n_K}|^2}\,\mathrm{d}V_{n_K} \le C_2^{n_K}n_K^{2n_K}.$$

Integrating the collision times

$$\int_{\mathfrak{T}_{\underline{n}_K}}\mathrm{d}\tau_{[n_K-1]} \le \prod_{k=0}^{K-1}\frac{\theta^{n_k-n_{k+1}}}{(n_k-n_{k+1})!} \le \frac{((K-1)\theta)^{n_{K-1}-1}}{(n_{K-1}-1)!}\frac{\theta^{n_K-n_{K+1}}}{(n_K-n_{K+1})!}$$
$$\le \frac{2^{n_K-1}t^{n_{K-1}-1}\theta^{n_K-n_{K+1}}}{(n_K-1)!}.$$

Finally, we multiply the two previous inequalities and $\frac{1}{(n_K-1)!}$. Using the Stirling formula, we obtain the expected estimation. □

Lemma 6.2. *Fix $\bar n := (n_1, \cdots, n_K)$. There exists a constant $\mathfrak{a} \in (0,1))$, depending only on the dimension, such that for any $\varepsilon > 0$ sufficiently small, we have*

$$\frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!}\sum_T \int_{\mathfrak{T}_{\underline{n}_K}\times\mathfrak{G}_T^0} \left|1 - \mathbb{1}_{\mathfrak{G}_T^\varepsilon}\right| \Lambda_T(V_{n_K}, \nu_{[n_K-1]})\,\mathrm{d}\nu_{[n_K-1]}\,\mathrm{d}\tau_{[n_K-1]}\;\mathrm{d}x_1 M^{\otimes n_K}\,\mathrm{d}V_{n_K} \leq C^{n_K} t^{n_K+10}\varepsilon^{\mathfrak{a}}. \tag{6.8}$$

This is an estimation of the set of parameters leading to a pathology (a recollision, a triple interaction, or an overlap). It is proven in Annex B.2. From Lemma 6.2 we deduce

$$|(6.5)| \leq C(Ct)^{n_K}\varepsilon^{\mathfrak{a}}\|g\|\,\|h\|.$$

Lemma 6.3. *Fix $\bar n := (n_1, \cdots, n_k)$, T, $\varepsilon > 0$ and $(x_1, (\tau_i, \nu_i)_i, V_{n_K}) \in \mathfrak{G}_T^\varepsilon$. We have*

$$\forall \tau \in [0,t],\ \left|\mathsf{X}^\varepsilon_{n_K}(\tau) - \mathsf{X}^0_{n_K}(\tau)\right| \leq 2n_K\mathbb{V}\sum_{i=1}^{n_K-1} \frac{\varepsilon\left|\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-)\right|}{\left|(\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-))\times\nu_i\right|^2}\,. \tag{6.9}$$

Proof. Thanks to the estimation of the interaction time (B.1), the i-th collision lasts at most a time $\frac{\varepsilon\left|\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-)\right|}{\left|(\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-))\times\nu_i\right|^2}$. Hence, the two trajectories $\mathsf{X}^\varepsilon_{n_K}(\tau)$ and $\mathsf{X}^0_{n_K}(\tau)$ have coincident velocities for τ outside the union of the interval

$$\bigcup_{i=1}^{n_K-1}\left[\tau_i, \tau_i + \frac{\varepsilon\left|\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-)\right|}{\left|(\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-))\times\nu_i\right|^2}\right].$$

During a collision, a particle can cross a distance smaller than $\frac{\varepsilon\mathbb{V}\left|\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-)\right|}{\left|(\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-))\times\nu_i\right|^2}$ which bounds the error that a collision creates. Hence, after n_K collisions, summing over all the possible particles, we obtain the expected bound.

□

Lemma 6.4. *Fix $\bar n := (n_1, \cdots, n_K)$. For any $\varepsilon > 0$ sufficiently small , we have*

$$|(6.4)| \leq C(Ct)^{n_K-1}\mathbb{V}\varepsilon^{1/2}\|g\|_1\|h\|_1. \tag{6.10}$$

Proof. We have forbid any recollision, multiple encounter , and overlap. Hence, the velocities of pseudo-trajectories of particles of sizes ε and 0 coincide. Using the inequality $|f(z) - f(z')| \leq (1 \wedge |z|)\|f\|_1$,

$$\begin{aligned}(6.4) \leq \frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!}\sum_T \int_{\mathfrak{T}_{\underline{n}_K}\times\mathfrak{G}_T^\varepsilon} 2n_K\mathbb{V}\|h\|_1\|g\|_1 \sum_{i=1}^{n_K-1} 1\wedge\frac{\varepsilon\left|\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-)\right|}{\left|(\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-))\times\nu_i\right|^2} \\ \times\Lambda(V_{n_K}, \nu_{[n_K-1]})\,\mathrm{d}\nu_{[n_K-1]}\,\mathrm{d}\tau_{[n_K-1]}\,\mathrm{d}x_1 M^{\otimes n_K}\,\mathrm{d}V_{n_K} \\ \leq \frac{\left(\frac{Ct}{\mathfrak{d}}\right)^{-n_K+1}\mathbb{V}\|h\|_1\|g\|_1}{(n_K!)^2}\sum_{1\leq i\leq n_K-1}\sum_T \int_{\mathfrak{G}_T^\varepsilon} 1\wedge\frac{\varepsilon\left|\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-)\right|}{\left|(\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-))\times\nu_i\right|^2}\Lambda_T(V_{n_K}, \nu_{[n_K-1]})\,\mathrm{d}\nu_{[n_K-1]}M^{\otimes n_K}\,\mathrm{d}V_{n_K}.\end{aligned}$$

We need to bound

$$\sum_T \int_{\mathfrak{G}_T^\varepsilon} 1\wedge\frac{\varepsilon\left|\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-)\right|}{\left|(\mathrm{v}_{q_i}(\tau_i^-) - \mathrm{v}_{q_i}(\tau_i^-))\times\nu_i\right|^2}\Lambda_T(V_{n_K}, \nu_{[n_K-1]})\,\mathrm{d}\nu_{[n_K-1]}M^{\otimes n_K}\,\mathrm{d}V_{n_K} \tag{6.11}$$

Note that $\mathrm{v}_q^0(\tau_i^+)$ does not depend on the $\tau_{[n_K-1]}$, but only on the order of the collisions.

Fix a collision tree $T = (q_i, q_i', s_i)_i$. We define for $i \in [1, n_K-1]$ the applications $\left(\Xi_T^i\right)_{1\leq i\leq n_K}$ as $\Xi_T^i = \mathrm{id}$ if $i = 1$, and

$$\Xi_T^i : (V_{n_K}, \nu_{[n_K-1]}) \mapsto \begin{cases}\left(v_1, \cdots, \overset{q_{i-1}}{v'_{q_{i-1}}}, \cdots, \overset{q'_{i-1}}{v'_{q'_{i-1}}}, \cdots, v_{n_K}, \nu_1, \cdots, \overset{i-1}{\nu'_{i-1}}, \cdots, \nu_{n_K-1}\right) \text{ if } s_i = 1 \\ \left(V_{n_K}, \nu_1, \cdots, \overset{i-1}{-\nu_{i-1}}, \cdots, \nu_{n_K-1}\right) \text{ if } s_i = 1\end{cases} \tag{6.12}$$

with the new velocities given by the scattering $(v'_{q_{i-1}}, v'_{q'_{i-1}}, \nu'_{i-1}) := \xi_\alpha((v'_{q_{i-1}}, v'_{q'_{i-1}}, \nu'_{i-1}))$. We have that

$$(\mathrm{V}_{n_K}(\tau_i^-), \nu^i_{[n_K-1]}) := \Xi_T^i\Xi_T^{i-1}\cdots\Xi_T^1(V_{n_K}, \nu_{[n_K-1]}).$$

Using that the Jacobian of the scattering ξ_α is 1 and the conservation by the scattering of the energy and angular momentum, the Jacobian of the transformation $\Xi_T^i \Xi_T^{i-1} \cdots \Xi_T^1$ is

$$\Lambda_T(V_{n_K}, \nu_{[n_K-1]})\, \mathrm{d}\nu_{[n_K-1]}\, \mathrm{d}x_1 M^{\otimes n_K}\, \mathrm{d}V_{n_K} \to \Lambda_T^{(i)}(V_{n_K}, \nu_{[n_K-1]})\, \mathrm{d}\nu_{[n_K-1]}\, \mathrm{d}x_1 M^{\otimes n_K}\, \mathrm{d}V_{n_K}$$

where we start now the velocity process at time τ_i^- with $\mathrm{V}_{n_K}(\tau_i^-) := V_{n_K}$ and

$$\Lambda_T^{(i)}(V_{n_K}, \nu_{[n_K-1]}) := \prod_{j=1}^{i-1} \left(\left(\mathsf{v}_{q_j}(\tau_j^+) - \mathsf{v}_{\bar{q}_j}(\tau_j^+) \right) \cdot \nu_j \right)_+ \prod_{j=i}^{n_k-1} \left(\left(\mathsf{v}_{q_j}(\tau_j^-) - \mathsf{v}_{\bar{q}_j}(\tau_j^-) \right) \cdot \nu_j \right)_- .$$

Hence,

$$(6.11) \le \sum_T \int_{\mathfrak{G}_T^\varepsilon} 1 \wedge \frac{\varepsilon |v_{q_i} - v_{q_i}|}{|(v_{q_i} - v_{q_i}) \times \nu_i|^2} \Lambda_T^{(i)}(V_{n_K}, \nu_{[n_K-1]})\, \mathrm{d}\nu_{[n_K-1]} M^{\otimes n_K}\, \mathrm{d}V_{n_K}.$$

Using the usual bound on $\sum_T \Lambda(V_{n_K}, \mathrm{d}\nu_{[n_K-1]})$ that can be adapted to Λ_i, and that for $\vec{e}_1 \in \mathbb{S}^{d-1}$,

$$\int_{\mathbb{S}^{d-1}} 1 \wedge \tfrac{\delta}{|\vec{e}_1 \times \sigma|^2} d\sigma \lesssim \int_0^\pi \left(1 \wedge \tfrac{\delta}{|\sin\theta|^2} \right) \sin^{d-2}\theta d\theta \lesssim \delta^{1/2},$$

one has

$$(6.11) \le \sum_{(q,q')} \int (C n_K |V_{n_K}|^2 + 1)^{n_K} 1 \wedge \tfrac{\varepsilon |v_q - v_{q'}|}{|(v_q - v_{q'}) \times \nu|^2} M^{\otimes n_K}\, \mathrm{d}V_{n_K}\, \mathrm{d}\nu$$

$$\le n_K^2 (C' n_K^2)^{n_K} \int \frac{\varepsilon^{1/2} e^{-\frac{|V_{n_K}|^2}{4}}\, \mathrm{d}V_{n_K}}{|v_1 - v_2|^{1/2}} \le \varepsilon^{1/2} (C'' n_K^2)^{n_K}$$

as $\frac{1}{|v_1 - v_2|^{1/2}}$ is an integrable singularity. This concludes the proof. □

Finally, we get for any h and g Lipschitz

$$\int \mu_\varepsilon^{n_K - 1}\, \Psi_{\underline{n}_K}^{0,t}[h]\, g_{n_K}^\varepsilon M^{\otimes n_K}\, \mathrm{d}Z_{n_K} = \frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!} \sum_T \sigma(T) \int_{\mathfrak{T}_{\underline{n}_K} \times \mathfrak{G}_T^0} h(\mathsf{z}_1^0(t,T) g_{n_K}(\mathsf{Z}_{n_K}^0(0,T))$$

$$\times \Lambda_T(V_{n_K}, \nu_{[n_K-1]})\, \mathrm{d}\nu_{[n_K-1]}\, \mathrm{d}\tau_{[n_K-1]}\ \mathrm{d}x_1 M^{\otimes n_K}\, \mathrm{d}V_{n_K}$$

$$+ O\left(\varepsilon^{\mathfrak{a}} (\tfrac{Ct}{\mathfrak{d}})^{n_K} \|h\|_1 \|g\|_1 \right).$$

and therefore

$$(6.13) \quad G_\varepsilon^{\mathrm{main}}(t) = \sum_{\substack{n_1 \le \cdots \le n_K \\ n_j - n_{j-1} \le 2^j}} \frac{\mathfrak{d}^{-n_K+1}}{(n_K-1)!} \sum_T \sigma(T) \int_{\mathfrak{T}_{\underline{n}_K} \times \mathfrak{G}_T^0} h(\mathsf{z}_1^0(t,T) g_{n_K}(\mathsf{Z}_{n_K}^0(0,T))$$

$$\times \Lambda_T(V_{n_K}, \nu_{[n_k-1]})\, \mathrm{d}\nu_{[n_k-1]}\, \mathrm{d}\tau_{[n_K-1]}\ \mathrm{d}x_1 M^{\otimes n_K}\, \mathrm{d}V_{n_K} + O\left(\varepsilon^{\mathfrak{a}} (\tfrac{Ct}{\mathfrak{d}})^{2^{K+1}} \|h\|_1 \|g\|_1 \right).$$

6.2. **Linearized Boltzmann equation.** We identify now the main part of (6.13).

Let $\mathbf{g}_\alpha$ be the solution of the linearized Boltzmann equation

$$\partial_t \mathbf{g}_\alpha(t) + v \cdot \nabla_x \mathbf{g}_\alpha(t) = \frac{1}{\mathfrak{d}} \mathscr{L}_\alpha \mathbf{g}_\alpha(t),$$
$$\mathbf{g}_\alpha(t=0) = g$$

where $\mathscr{L}_\alpha$ is the linearized Boltzmann operator associated to the potential $\alpha \mathscr{V}(\cdot)$

$$\mathscr{L}_\alpha g(v) := \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \big(g(v') + g(v_*') - g(v) - g(v_*) \big) M(v_*) \big((v - v_*) \cdot \nu \big)_+\, \mathrm{d}\nu\, \mathrm{d}v_*.$$

This equation can be rewritten in the Duhamel form:

$$\mathbf{g}_\alpha(t) = S(t) g + \frac{1}{\mathfrak{d}} \int_0^t S(t - \tau_1) \mathscr{L}_\alpha \mathbf{g}_\alpha(\tau_1)\, \mathrm{d}\tau_1$$

where $S(\tau)$ is the free transport

$$S(\tau) g(x,v) = g(x - tv, v).$$

We iterate this formula, but we still want to cut the cases with too many collisions in a short time interval (as for the particle system). Let's define

$$Q_{m,n}(\tau)[g] = \frac{1}{\mathfrak{d}^{m-n}} \int_0^\tau \mathrm{d}\tau_n \int_0^{\tau_n} \cdots \int_0^{t_{m+2}} \mathrm{d}\tau_{m+1} S(t-\tau_n) \mathscr{L}_\alpha S(\tau_n - \tau_{n-1}) \cdots \mathscr{L}_\alpha S(\tau_{m+1}) g,$$

and for $\underline{n}_k := (n_1, \cdots, n_k)$ with $1 \le n_1 \le \cdots \le n_k$,

$$Q_{\underline{n}_k}(\tau) g = Q_{1,n_1}(\tfrac{\tau}{k}) Q_{n_1,n_2}(\tfrac{\tau}{k}) \cdots Q_{n_{k-1},n_k}(\tfrac{\tau}{k})[g].$$

We have

$$\mathbf{g}(t) = \sum_{\substack{n_1 \le \cdots \le n_K \\ n_j - n_{j-1} \le 2^j}} Q_{\underline{n}_k}(t)[g] + \sum_{k=1}^{K} \sum_{\substack{n_1 \le \cdots \le n_{k-1} \\ n_j - n_{j-1} \le 2^j}} \sum_{n_k > +n_{k-1}+2^k} Q_{\underline{n}_k}(k\theta)[\mathbf{g}_\alpha(t-k\theta)]. \tag{6.14}$$

In a first time, we bound the term of the sum: we have the classical estimates

Proposition 6.5. *There exists a constant C such that for any $g \in L^2(M(v)dz)$, and $\underline{n} := (n_1, \cdots, n_k)$,*

$$\left\| Q_{\underline{n}}(k\theta) g \right\|_{L^2(M^2(v)dz)} \le \left(\tfrac{C(k-1)\theta}{\mathfrak{d}}\right)^{\frac{n_{k-1}}{2}} \left(\tfrac{C\theta}{\mathfrak{d}}\right)^{\frac{n_k - n_{k-1}}{2}} \|g\|_{L^2(M(v)dz)}. \tag{6.15}$$

The proof is the same as the one of Proposition 7.5 of [LB25].

Because $\mathbf{g}_\alpha(t)$ is bounded in $L^\infty_t L^2(M(v)dz)$ by $\|g\|_{L^2(M(v)dz)} \le C\|g\|_0$, we can bound the rest term of (6.14) by

$$\begin{aligned}
&\left| \sum_{k=1}^{K} \sum_{\substack{n_1 \le \cdots \le n_{k-1} \\ n_j - n_{j-1} \le 2^j}} \sum_{n_k > +n_{k-1}+2^k} \int h(z) Q_{\underline{n}_k}(k\theta)[\mathbf{g}_\alpha(t-k\theta)](z) M(v)\,\mathrm{d}z \right| \\
&\qquad \le \sum_{k=1}^{K} \sum_{\substack{n_1 \le \cdots \le n_{k-1} \\ n_j - n_{j-1} \le 2^j}} \sum_{n_k > +n_{k-1}+2^k} \left(\tfrac{C(k-1)\theta}{\mathfrak{d}}\right)^{\frac{n_{k-1}}{2}} \left(\tfrac{C\theta}{\mathfrak{d}}\right)^{\frac{n_k - n_{k-1}}{2}} \|g\|_0 \|h\|_0 \\
&\qquad \le \sum_{k=1}^{K} \sum_{\substack{n_1 \le \cdots \le n_{k-1} \\ n_j - n_{j-1} \le 2^j}} \sum_{n_k > 2^k + n_{k-1}} \left(C\tfrac{\theta t}{\mathfrak{d}^2}\right)^{\frac{n_k - n_{k-1}}{2}} \|g\|_0 \|h\|_0 \\
&\qquad \le C \sum_{k=1}^{K} 2^{k^2} \left(C\tfrac{\theta t}{\mathfrak{d}^2}\right)^{2^{k-1}} \|g\| \|h\| \le C\tfrac{\theta t}{\mathfrak{d}^2} \|g\|_0 \|h\|_0.
\end{aligned} \tag{6.16}$$

The series converges since $\frac{\theta t}{\mathfrak{d}^2} < 1$.

The final step is the identification of the main part in (6.14):

Proposition 6.6. *Fix $\underline{n}_K := (n_1, \cdots, n_K)$ an increasing sequence of integer. Then*

$$\begin{aligned}
\int_{\mathbb{D}} h(z) Q_{\underline{n}_k}(t)[g](z) M(v) dz = \frac{\mathfrak{d}^{-n_k+1}}{(n_k-1)!} \sum_T \sigma(T) \int_{\mathfrak{T}_{\underline{n}_K} \times \mathfrak{G}^0_T} & h(\mathsf{z}^0_1(t,T) g_{n_k}(\mathsf{Z}^0_{n_k}(0,T)) \\
& \times \Lambda_T(V_{n_K}, \nu_{[n_k-1]})\, \mathrm{d}\nu_{[n_k-1]}\, \mathrm{d}\tau_{[n_k-1]}\, \mathrm{d}x_1 M^{\otimes n_k}\, \mathrm{d}V_{n_k}.
\end{aligned} \tag{6.17}$$

Remark 6.1. *It is the counterpart of the Step 2 of the proof of Proposition IV.1 of* [BGSRS22].

Proof. We fix for the moment the collision times $(\tau_i)_i$.

Definition 6.2. *Fix a collision tree $T := (q_i, \bar{q}_i, \bar{s}_i)_{i \le n_k - 1)}$ and a final particle q_f. We say that a sequence $(i_1, \cdots, i_\ell)$ is* causal *if*

$$i_1 < \cdots < i_\ell, \ \forall j < \ell, \ \{q_{i_j}, \bar{q}_{i_j}\} \cap \{q_{i_{j+1}}, \bar{q}_{i_{j+1}}\} \neq \emptyset.$$

A particle $\bar{q}$ influences *the particle 1 (respectively q_f) if there exists a causal sequence $(i_1, \cdots, i_\ell)$ such that $\bar{q} \in \{q_{i_1}, \bar{q}_{i_1}\}$ and $1 \in \{q_{i_\ell}, \bar{q}_{i_\ell}\}$ (respectively $q_f \in \{q_{i_1}, \bar{q}_{i_1}\}$ and $\bar{q} \in \{q_{i_\ell}, \bar{q}_{i_\ell}\}$).*

We begin by developing g_{n_K}:

$$\sum_T \sigma(T) \int_{\mathbb{G}_T^0} h(\mathsf{z}_1^0(t,T)) g_{n_k}(\mathsf{Z}_{n_k}^0(0,T)) \Lambda_T(V_{n_K}, \nu_{[n_k-1]}) \, \mathrm{d}\nu_{[n_k-1]} \, \mathrm{d}\tau_{[n_k-1]} \, \mathrm{d}x_1 M^{\otimes n_k} \, \mathrm{d}V_{n_k}$$
$$= \sum_{q_f=1}^{n_k} \sum_T \sigma(T) \int_{\mathbb{G}_T^0} h(\mathsf{z}_1^0(t,T)) g(\mathsf{z}_{q_f}^0(0,T)) \Lambda_T(V_{n_K}, \nu_{[n_k-1]}) \, \mathrm{d}\nu_{[n_k-1]} \, \mathrm{d}\tau_{[n_k-1]} \, \mathrm{d}x_1 M^{\otimes n_k} \, \mathrm{d}V_{n_k}. \tag{6.18}$$

Suppose that there exists a particle $\bar{q}$ that does not influence 1. If $\bar{q}$ has more than two neighbor particles, one of them is farther from the particle 1 in the graph T (the graph T is a tree). Then, this new particle does not influence 1. We deduce that without loss of generality, we can suppose that q has only one neighbor. Then, it has only one collision $\bar{\iota}$.

We now use the application Ξ_T^i defined in (6.12). We recall that

$$\Xi_T^{\bar{\iota}} \Xi_T^{\bar{\iota}-1} \cdots \Xi_T^1 (V_{n_K}, \nu_{[n_K-1]}) = (\tilde{V}_{n_K} = \mathsf{V}_{n_K}(\tau_{\bar{\iota}}^-), \tilde{\nu}_{[n_K-1]}).$$

In a second time, for a fixed $(V_{n_K}, \nu_{[n_K-1]})$, we perform the translation $x_1 \mapsto \tilde{x}_1 := \mathsf{x}_1(\tau_{\bar{\iota}})$. The Jacobian of $\tau_{\mathsf{x}_i(\tau_i)} \Xi_T^{\bar{\iota}} \Xi_T^{\bar{\iota}-1} \cdots \Xi_T^1$ is

$$\Lambda_T(V_{n_K}, \nu_{[n_K-1]}) \, \mathrm{d}\nu_{[n_K-1]} \, \mathrm{d}x_1 M^{\otimes n_K} \, \mathrm{d}V_{n_K} \to \Lambda_T^{(i)}(\tilde{V}_{n_K}, \nu_{[n_K-1]}) \, \mathrm{d}\nu_{[n_K-1]} \, \mathrm{d}\tilde{x}_1 M^{\otimes n_K} \, \mathrm{d}\tilde{V}_{n_K}.$$

We start now the velocity process at time τ_i^- with $\mathrm{V}_{n_K}(\tau_i^-) := \tilde{V}_{n_K}$ and

$$\Lambda_T^{(i)}(\tilde{V}_{n_K}, \nu_{[n_K-1]}) := \prod_{j=1}^{i-1} \left(\left(\mathsf{v}_{q_j}(\tau_j^+) - \mathsf{v}_{\bar{q}_j}(\tau_j^+) \right) \cdot \nu_j \right)_+ \prod_{j=i}^{n_k-1} \left(\left(\mathsf{v}_{q_j}(\tau_j^-) - \mathsf{v}_{\bar{q}_j}(\tau_j^-) \right) \cdot \nu_j \right)_- .$$

We pair T with the tree $\tilde{T}$ as

$$\tilde{T} := \begin{cases} (q_j, \bar{q}_j, \bar{s}_j) & \text{for } j \neq i \\ (q_j, \bar{q}_j, -\bar{s}_j) & \text{for } j = i. \end{cases}$$

Then $\sigma(T) = -\sigma(\tilde{T})$, and for same $(\tilde{V}_{n_K}, \tau_{[n_K-1]}, \nu_{[n_K-1]})$, we have $\mathsf{z}_1^0(t,T) = \mathsf{z}_1^0(t,\tilde{T})$ and $\mathsf{z}_q^0(0,T) = \mathsf{z}_q^0(0,\tilde{T})$. We have $\Lambda_T^{(i)}(V_{n_K}, \nu_{[n_K-1]}) = \Lambda_{\tilde{T}}^{(i)}(V_{n_K}, \nu_{[n_K-1]})$. Thus

$$\sigma(T) \int_{\mathbb{G}_T^0} h(\mathsf{z}_1^0(t,T)) g(\mathsf{z}_q^0(0,T)) \Lambda(V_{n_K}, \nu_{[n_K-1]}) \, \mathrm{d}\nu_{[n_K-1]} \, \mathrm{d}x_1 M^{\otimes n_K} \, \mathrm{d}V_{n_K}$$
$$= -\sigma(\tilde{T}) \int_{\mathbb{G}_{\tilde{T}}^0} h(\mathsf{z}_1^0(t,\tilde{T})) g(\mathsf{z}_q^0(0,\tilde{T})) \Lambda(V_{n_K}, \nu_{[n_K-1]}) \, \mathrm{d}\nu_{[n_K-1]} \, \mathrm{d}x_1 M^{\otimes n_K} \, \mathrm{d}V_{n_K}.$$

The same strategy can be reproduced if there exists some particle that does not influence the final particle q_f. Hence, it remains only in (6.18) the trees such that every particle influences both 1 and q_f. The other terms are exactly compensated.

For a remaining tree $T = (q_i, \bar{q}_i, \nu_i)_i$ we can prove the following lemma:

Lemma 6.7. *For all i, the set $\{q_i, \bar{q}_i\} \cap \{q_{i+1}, \bar{q}_{i+1}\}$ has exactly one element.*

Proof. First, there exists a causal sequence $(q_{i_1}, \bar{q}_{i_1}), \cdots, (q_{i_k}, \bar{q}_{i_k})$ such that $q_f \in \{q_{i_1}, \bar{q}_{i_1}\}$ and $1 \in \{q_{i_k}, \bar{q}_{i_k}\}$.

Consider $q \notin \bigcup_l \{(q_{i_k}, \bar{q}_{i_k})\}$. There exist two causal paths $(q_{j_1}, \bar{q}_{j_1}), \cdots, (q_{j_\ell}, \bar{q}_{j_\ell})$ from q_f to q and $(q_{j'_1}, \bar{q}_{j'_1}), \cdots, (q_{j'_{\ell'}}, \bar{q}_{j'_{\ell'}})$ from q to 1. We define ℓ_0 such that $\forall l \leq \ell_0$, $(q_{j_{\ell-l+1}}, \bar{q}_{j_{\ell-l+1}}) = (q_{j'_l}, \bar{q}_{j'_l})$. The sequence $(q_{j'_1}, \bar{q}_{j'_1}), \cdots, (q_{j'_{\ell_0}}, \bar{q}_{j'_{\ell_0}})$ is both increasing and decreasing, thus $\ell_0 \leq 1$. The sequence $(q_{j'_1}, \bar{q}_{j'_1}), \cdots, (q_{j_{\ell-\ell_0}}, \bar{q}_{j_{\ell-\ell_0}}), (q_{j'_{\ell_0+1}}, \bar{q}_{j'_{\ell_0+1}}), \cdots, (q_{j'_{\ell'}}, \bar{q}_{j'_{\ell'}})$ is a causal path from q_f to 1. We deduce that $\ell_0 = 1$.

Finally, for any $l \leq k-1$, $j \in]i_l, i_{l+1}[$, one $\{q_j, \bar{q}_j\} \cap \{q_{i_l}, \bar{q}_{i_l}\} \cap \{q_{i_{l+1}}, \bar{q}_{i_{l+1}}\}$. The conclusion follows □

Using the lemma, we can construct the sequences $(\tilde{q}_i)_{0 \leq i \leq n_K-1}$, $(\tilde{q}'_i)_{1 \leq i \leq n_K-1}$ and $(\tilde{s}_i)_{1 \leq i \leq n_K-1}$ by

$$\tilde{q}_0 := q_f, \ \tilde{q}_{n_K-1} := 1, \ \{\tilde{q}_i\} := \{q_i, \bar{q}_i\} \cap \{q_{i+1}, \bar{q}_{i+1}\}, \ \{\tilde{q}'_i\} := \{q_i, \bar{q}_i\} \setminus \{\tilde{q}_i\}$$

and

$$\tilde{s}_i := \begin{cases} 1 \text{ if } q_i = q_{i-1} \\ -1 \text{ else.} \end{cases}$$

The sequence $(\tilde{q}'_i)_i$ encodes the order in which particles collide. In addition, we can reconstruct T for a given sequence $(\bar{s}_i, \tilde{s}_i, \tilde{q}'_i)_i$.

We reorder the particles such that $\tilde{q}'_i = n_K - i$ (there are $(n_K-1)!$ possibilities).

Finally, we have to identify the four possible $(\bar{s}_i, \tilde{s}_i)_i$ with the four parts of $\mathscr{L}_\alpha$: $(1,1)$ with $g(v')$ (we follow the same particle that is deviated by the collision), $(1,-1)$ with $-g(v)$ (we follow the same particle that is not deviated by the collision), $(-1,1)$ with $g(v'_*)$ and $(-1,-1)$ with $-g(v_*)$. There are $(n_K-1)!$ possible sequences $(\tilde{q}'_i)_i$.

We conclude that

$$\frac{1}{(n_K-1)!}\sum_T \sigma(T)\int_{\mathbb{G}^0_T} h(\mathsf{z}^0_1(t,T))g_{n_K}(\mathsf{Z}^0_{n_K}(0,T))\Lambda(V_{n_K},\nu_{[n_K-1]})\,\mathrm{d}\nu_{[n_K-1]}\,\mathrm{d}x_1 M^{\otimes n_K}\,\mathrm{d}V_{n_K}$$
$$= \int_{\mathbb{D}} h(z)S(t-\tau_{n_K-1})\mathscr{L}_\alpha S(\tau_{n_K-1}-\tau_{n_K-2})\cdots\mathscr{L}_\alpha S(\tau_1)g(z)\ M(v)\,\mathrm{d}z.$$

We obtain the expected result by integrating with respect to $(\tau_1,\cdots,\tau_{n_K-1})$. □

Combining the preceding proposition and the estimations (6.16) and (6.13), we obtain:

$$G^{\mathrm{main}}_\varepsilon(t) = \int_{\mathbb{D}} h(z)\mathbf{g}_\alpha(t,z)M(z)\,\mathrm{d}z + O\left(\left(\tfrac{\theta t}{\mathfrak{d}^2}+\varepsilon^{\mathfrak{a}}K2^{K^2}(\tfrac{Ct}{\mathfrak{d}})^{2^{K+1}}\right)\|h\|_1\|g\|_1\right). \tag{6.19}$$

7. Estimation of non-pathological recollisions

In the last two sections, we estimate the error terms where the pseudotrajectory can have a recollision. We begin with the case of non-pathological recollision.

$$G^{\mathrm{rec},1}_\varepsilon(t) = \sum_{\substack{0\le k\le K-1\\ 1\le k'\le K'}}\ \sum_{\substack{(n_j)_{j\le k}\\ 0\le n_j-n_{j-1}\le 2^j}}\ \sum_{n_{k+2}\ge n_{k+2}\ge n_k} \mathbb{E}_\varepsilon\left[\frac{1}{\sqrt{\mu}}\sum_{\underline{i}_{n_k}}\Psi^{>,t-t_s}_{\underline{n}_{k+1}}[h]\left(\mathbf{Z}_{\underline{i}_{n_k}}(t_s)\right)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]$$

where $t-t_s = k\theta + k'\theta'$.

Proposition 7.1. *There exists a constant $\mathfrak{a}\in(0,1))$, depending only on the dimension, such that for ε small enough,*

$$\left|G^{rec,1}_\varepsilon(t)\right| \le \|g\|\|h\|\varepsilon^{\mathfrak{a}/2}(C't)^{2^{t/\theta}+d+9}. \tag{7.1}$$

It is sufficient to prove the two following estimations:

Proposition 7.2. *Fix $k\in\mathbb{N}$, $\underline{n} := (n_1,\cdots,n_{k+2})\in\mathbb{N}^k$. Then fixing $x_1=0$ we have*

$$\int \sup_{y\in\mathbb{T}}\left|\Psi^{>,t-t_s}_{\underline{n}_{k+2}}[h](\mathrm{tr}_y\, Z_{n_{k+2}})\right|\frac{e^{-\mathscr{H}_{n_{k+2}}}}{(2\pi)^{\frac{n_k d}{d}}}\,\mathrm{d}V_{n_{k+2}}\,\mathrm{d}X_{2,n_{k+2}}$$
$$\le \varepsilon^{\mathfrak{a}}\frac{\|h\|_0}{(\mu\mathfrak{d})^{n_{k+2}-1}}C^{n_{k+2}}\delta^2\theta^{(n_{k+2}-n_k-3)_+}t^{n_k+9+d}\varepsilon^{\mathfrak{a}}, \tag{7.2}$$

and, for $m\in[1,n_{k+2}]$,

$$\int \sup_{y\in\mathbb{T}}\left|\Psi^{>,t-t_s}_{\underline{n}_{k+2}}[h]\circledast_l\Psi^{>,t-t_s}_{\underline{n}_{k+2}}[h](\mathrm{tr}_y\, Z_{2n_{k+2}-m})\right|M^{\otimes(2n_{k+2}-m)}\,\mathrm{d}V_{2n_{n+2}-m}\,\mathrm{d}X_{2,2n_{k+2}-m}$$
$$\le \frac{\mu^{m-1}}{n^m_{k+2}}\varepsilon^{\mathfrak{a}}\left(\frac{\|h\|_0}{(\mu\mathfrak{d})^{n_{k+2}-1}}C^{n_{k+2}}\right)^2\delta^2\theta^{(n_{k+2}-n_k-3)_+}t^{n_k+n_{k+2}+9+d}. \tag{7.3}$$

Using these estimations and Corollary 4.3,

$$\left|\mathbb{E}_\varepsilon\left[\mu^{-\frac12}\sum_{\underline{i}_{n_{k+2}}}\Psi^{>,t-t_s}_{\underline{n}_{k+2}}[h]\left(\mathbf{Z}_{\underline{i}_{n_{k+2}}}(t_s)\right)\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]\right|$$
$$\le \|h\|_0\|g\|_0\frac{C^{n_{k+2}}}{\mathfrak{d}^{n_{k+2}-1}}\left(\varepsilon^{\frac12+\mathfrak{a}}\theta^{(n_{k+2}-n_k-3)_+}\delta^2t^{n_k+d+9}+\left(e^{\mathfrak{a}}\theta^{(n_{k+2}-n_k-3)_+}\delta^2t^{n_k+d+9+m}\right)^{\frac12}\right)$$
$$\le \|g\|_0\|h\|_0\delta\varepsilon^{\frac{\mathfrak{a}}{2}}C^{n_{k+2}}(\tfrac{\theta}{\mathfrak{d}})^{\frac{(n_{k+2}-n_k-3)_+}{2}}(\tfrac{t}{\mathfrak{d}})^{\frac{n_{k+2}+n_k}{2}+d+9}$$
$$\le \|g\|_0\|h\|_0\delta\varepsilon^{\frac{\mathfrak{a}}{2}}(\tfrac{Ct}{\mathfrak{d}})^{n_k+d+9}(\tfrac{Ct\theta}{\mathfrak{d}^2})^{\frac{(n_{k+2}-n_k-3)_+}{2}}.$$

Using that $\frac{Ct\theta}{\mathfrak{d}^2} < 1$ and $K'\delta = \theta \leq 1$, we can sum on k, k' and $\underline{n}_{k+2}$

$$\begin{aligned}\left|G_\varepsilon^{\mathrm{rec},1}(t)\right| &\leq \sum_{\substack{1\leq k\leq K-1\\ 1\leq k'\leq K'}} \sum_{n_1\leq\cdots\leq n_k} \sum_{n_{k+2}\geq n_{k+1}\geq n_k} \|g\|_0\|h\|_0\delta\varepsilon^{\frac{\mathfrak{a}}{2}}(\tfrac{Ct}{\mathfrak{d}})^{n_k+d+9}(\tfrac{Ct\theta}{\mathfrak{d}^2})^{\frac{(n_{k+2}-n_k-2)_+}{2}}\\ &\leq \|g\|_0\|h\|_0 K'\delta\varepsilon^{\frac{\mathfrak{a}}{2}} K^{K^2}(\tfrac{Ct}{\mathfrak{d}})^{2^K+d+9}\\ &\leq \|g\|_0\|h\|_0\varepsilon^{\frac{\mathfrak{a}}{2}}(\tfrac{C't}{2\mathfrak{d}})^{2^K+d+9}\end{aligned}$$

This concludes the proof of (7.1).

Proof of (7.2). We recall that the pseudotrajectory development takes the form

$$\Psi^{>,t-t_s}_{\underline{n}_{k+2}}[h] := \frac{1}{(n_{k+2}-1)!}\sum_{\substack{1\in\omega\subset[n_{k+2}]\\ |\omega|=n_{k+1}\\ (s_i)_{i\leq n_{i+2}-1}}}\prod_{k=1}^{n_l-1} s_k\, h(\mathsf{z}_q(t,\cdot,\{q\},\omega,(s_i)_i))\mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}}\prod_{i=1}^{k}\mathbb{1}_{\mathfrak{n}(t-i\theta)=n_i}.$$

Here $\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}$ is the set of initial configurations $Z_{n_{k+2}}$ such that the pseudotrajectory has

- 1 the final particle at time $t-t_s$,
- ω the set of particles at time δ,
- at least one recollision,
- no pathological recollision (thanks to the conditioning on $\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}$).

Lemma 7.3. *There exists a constant $\mathfrak{a}\in(0,1)$ such that for any $\underline{n}$, k' and $(s_i)_i$,*

$$\int \mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}} M^{\otimes n_{k+2}}\,\mathrm{d}X_{2,n_{k+2}}\,\mathrm{d}V_{n_{k+2}} \leq C'\left(\frac{C'n_{k+2}}{\mu\mathfrak{d}}\right)^{n_{k+2}-1}\delta^2\theta^{(n_{k+2}-n_k-2)_+}t^{n_k+2d+4}\varepsilon^{\mathfrak{a}}. \tag{7.4}$$

Proof. We may define the clustering tree $T^>$ as before, by looking at collisions in temporal order and keeping only the clustering collisions. However, this will not be sufficient to characterize the initial data.

Let $(\bar q,\bar q')$ (with $\bar q<\bar q'$) be the first two particles having a non-clustering collision, τ_{cycle} the time of this collision, and $c\in[1,n_{k+2}-1]$ such that τ_{cycle} lies between the times of the c-th and the $(c+1)$-th clustering collision. The parameters $\bar T := (T^>,(\bar q,\bar q',c))$ provide a partition of the set of initial data.

We denote

$$\mathfrak{T}_{\underline{n}_{k+2}} := \left\{\begin{array}{c}(\tau_i)_{i\leq n_{k+2}-1},\ \tau_i\leq\tau_{i+1},\\ \forall j\leq n_{k+2}-n_{k+1},\ \tau_j\leq\delta\\ \forall j\leq n_{k+2}-n_k[,\ \tau_j\leq k'\delta\\ \forall\ell\leq k,\ j\leq n_{k+2}-n_{k+2-\ell-1},\ \tau_j\leq k'\delta+(\ell+1)\theta\end{array}\right\}$$

For a given initial data Z_n and $\bar T := (T^>,(\bar q,\bar q',c))$, we define τ_i as the time of the i-th clustering collision and $\nu_i := (\mathrm{x}_{q_i}(\tau_i)-\mathrm{x}_{q_i}(\tau_i'))/\varepsilon$. We denote $\mathfrak{T}_{\underline{n}_{k+2}}\times\mathfrak{G}^{>,t-t_s,\bar T}_{\{q\},\omega,(s_i)_i}$ the image of the set of initial datum

$$\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}\cap\{T^>\text{ is the collision tree, first collision implies }(\bar q,\bar q')\text{ during }(\tau_c,\tau_{c+1})\}$$

by the application $(X_{2,n_{k+2}},V_{n_{k+2}})\to(\tau_{[n_{k+2}-1]},\nu_{[n_{k+2}-1]},V_{n_{k+2}})$.

$$\begin{aligned}&\int \mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}}\frac{e^{-\mathscr{H}_{n_{k+2}}}}{(2\pi)^{\frac{n_{k+2}d}{2}}}\,\mathrm{d}X_{2,n_{k+2}}dV_{n_{k+2}}\\ &\qquad=\frac{1}{(\mu\mathfrak{d})^{n_{k+2}-1}}\sum_{\bar T}\int_{\mathfrak{T}_{\underline{n}_{k+2}}\times\mathfrak{G}^{>,t-t_s,\bar T}_{\{q\},\omega,(s_i)_i}}\prod_{i=1}^{n_{k+2}-1}|(\mathrm{v}^\varepsilon_{q_i}(\tau_i)-\mathrm{v}^\varepsilon_{q_i'}(\tau_i))\cdot\nu_i|\,\mathrm{d}\nu_i\,\mathrm{d}\tau_i M^{\otimes n_{k+2}}dV_{n_{k+2}}.\end{aligned}$$

Using Lemma B.2, we get

$$\begin{aligned}\int \mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}}\frac{e^{-\mathscr{H}_{n_{k+2}}}}{(2\pi)^{\frac{n_{k+2}d}{2}}}\,\mathrm{d}X_{2,n_{k+2}}dV_{n_{k+2}} &\leq\frac{(Cn_{k+2})^{n_{k+2}}}{(\mu\mathfrak{d})^{n_{k+2}-1}}t^{n_k-1}\theta^{(n_{k+2}-n_k-1)_+}\varepsilon^{1/4}\\ &\leq\frac{(Cn_{k+2})^{n_{k+2}}}{(\mu\mathfrak{d})^{n_{k+2}-1}}t^{n_k-1}\theta^{(n_{k+2}-n_k-1)_+}\delta^2\varepsilon^{1/12}\end{aligned} \tag{7.5}$$

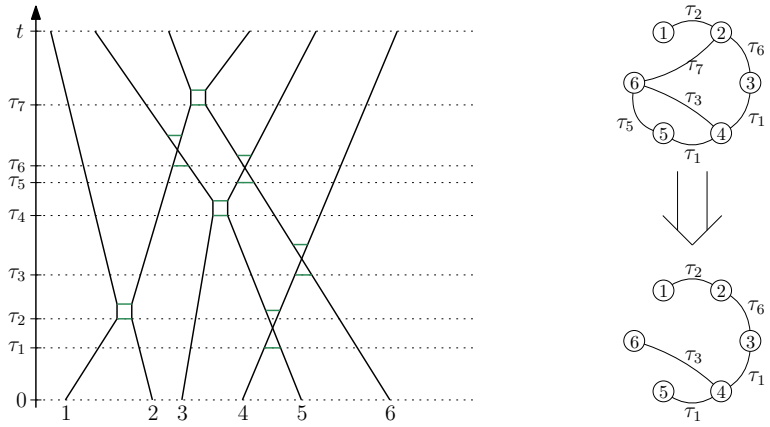

FIGURE 9. Example of construction of a clustering tree.

using that $\delta = \varepsilon^{1/12}$.

□

We obtain the expected result by summing on

$$(s_i)_{i\le n_{k+2}-1} \in \{\pm 1\}^{n_{k+2}-1},\ \omega \subset [n_{k+2}],\ q \in \omega$$

and dividing by $n_{k+2}!$. □

Proof of (7.3). We use first the same bound as in the proof of (7.2) and of (5.3),

$$\begin{aligned}(7.6)\quad &\left|\Psi^{>,t-t_s}_{\underline{n}_{k+2}}[h] \circledast_m \Psi^{>,t-t_s}_{\underline{n}_{k+2}}[h](Z_{2n_{k+2}-m})\right| \\ &\le \frac{\|h\|^2}{(n_{k+2}!)^2}\frac{(n_{k+2}-m)!^2 m!}{n_{k+2}^2(2n_{k+2}-m)!}\sum_{\substack{\bar\omega\cup\bar\omega'=[2n_{k+2}-m]\\ |\bar\omega|=|\bar\omega'|=n_{k+2}}}\ \sum_{\substack{q\in\omega\subset\bar\omega\\ |\omega|=n_{k+1}\\ (s_i)_{i\le n_{i+2}-1}}}\ \sum_{\substack{q'\in\omega'\subset\bar\omega'\\ |\omega'|=n_{k+1}\\ (s'_i)_{i\le n_{i+2}-1}}} \mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}}(Z_{\bar\omega}) \\ &\qquad\qquad \times \mathbb{1}_{\mathfrak{n}(k'\delta)=n_k}\mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q'\},\omega',(s'_i)_i}}(Z_{\bar\omega'}).\end{aligned}$$

where $\mathfrak{n}(\theta)$ is the number of particles at time θ in the pseudotrajectory $\mathrm{Z}(\tau)$. Note that the formula is invariant under translation. We can then set $x_1 = 0$ and integrate with respect to the other variables.

Using the same strategy as in the proof of (5.3), we have

$$\int \mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q'\},\omega',(s'_i)_i}}(Z_{\bar\omega'})e^{-\frac12\mathscr{H}_{2n_{k+2}-m}}\,\mathrm{d}Z_{\bar\omega'\setminus\bar\omega} \le C\left(\frac{C}{\mu\mathfrak{d}}\right)^{n_k-m} t^{n_{k+2}-m}(2n_{k+2}-m)^{n_{k+2}-m}.$$

The sum over the remaining particles is estimated using (7.5)

$$\begin{aligned}\int \mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q'\},\omega',(s'_i)_i}}(Z_{\bar\omega'})\mathbb{1}_{\mathscr{R}^{>,t-t_s}_{\{q\},\omega,(s_i)_i}}(Z_\omega)\frac{e^{-\mathscr{H}_{2n_{k+2}-m}}}{(2\pi)^{\frac{(2n_{k+2}-m)d}{2}}}\,\mathrm{d}X_{1,2n_{k+2}-m}dV_{2n_{k+2}-m}& \\ \le \frac{(C2n_{k+2})^{2n_{k+2}-m}}{(\mu\mathfrak{d})^{2n_{k+2}-m-1}}\delta^2\theta^{(n_{k+2}-n_k-3)_+}t^{n_k+n_{k+2}+9+d}\varepsilon^{\mathfrak{a}}.&\end{aligned}$$

We obtain the expected result by combining the two estimations, summing on the possible parameters $((s_i)_i, \bar\omega, \omega, q)$ and $((s'_i)_i, \bar\omega', \omega', q')$ and then dividing by $(n_{k+2})!^2$. □

8. Estimation of the local recollisions

In the present section, we discuss $G^{\mathrm{rec},2}_\varepsilon(t)$ defined by

$$\begin{aligned}G^{\mathrm{rec},2}_\varepsilon(t) = \sum_{\substack{0\le k\le K-1\\ 1\le k'\le K'}}\ \sum_{\substack{(n_j)_{j\le k}\\ 0\le n_j-n_{j-1}\le 2^j}}\ \sum_{n_{k+2}\ge n_{k+1}\ge n_k} \mathbb{E}_\varepsilon\Bigg[\frac{1}{\sqrt{\mu}}\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}& \\ \times \sum_{\underline{i}_{n_{k+2}}} \Phi^{>,\delta}_{n_{k+1},n_{k+2}}\Psi^{,t-t_s-\delta}_{\underline{n}_{k+1}}[h]\left(\mathbf{Z}_{\underline{i}_{n_{k+2}}}(t_s)\right)\Bigg]&\end{aligned}$$

for $t_s := t - k\theta - k'\delta$.

We will prove the following bound:

Proposition 8.1. *There exists a constant $\mathfrak{a} \in (0,1)$), depending only on the dimension, such that for $\varepsilon > 0$ small enough, we have*

$$\left|G_\varepsilon^{rec,2}(t)\right| \leq C\|h\|_0\|g\|_0(C\tfrac{t}{\mathfrak{d}})^{2^{K+1}}\varepsilon^{\frac{\mathfrak{a}}{2}}. \tag{8.1}$$

In the following, we denote

$$\bar\Phi^{k'}_{\underline{n}_{k+2}}(Z_{n_{k+2}}) := \frac{1}{(n_{k+2})!}\sum_{\sigma\in\mathfrak{S}_{n_{k+2}}} \Phi^{>,\delta}_{n_{k+1},n_{k+2}}\Psi^{,t-t_s-\delta}_{\underline{n}_{k+1}}[h]\left(Z_{\sigma[n_{k+2}]}\right)$$

The aim of this part is to prove the following bound on $\Phi^{k'}_{\underline{n}_{k+2},p}$:

Proposition 8.2. *Fix $n_1 \leq \cdots \leq n_{k+2} \leq p$. For $m \in \{1,\cdots p\}$ we have for $x_1 = 0$*

$$\int \sup_{y\in\mathbb{T}}\left|\bar\Phi^{k'}_{\underline{n}_{k+2}}(\mathrm{tr}_y\, Z_{n_{k+2}})\right|\frac{e^{-\mathscr{H}_{n_{k+2}}(Z_{n_{k+2}})}}{(2\pi)^{\frac{dn_{k+2}}{2}}}\,\mathrm{d}Z_{2,n_{k+2}}\,\mathrm{d}v_1 \leq \frac{\|h\|_0}{(\mu\mathfrak{d})^{n_{k+2}-1}}C^{n_{k+2}}\delta^2\varepsilon^{\mathfrak{a}}\theta^{(n_{k+2}-n_k-2)_+}t^{n_k-1}, \tag{8.2}$$

$$\int \sup_{y\in\mathbb{T}}\left|\bar\Phi^{k'}_{\underline{n}_{k+2}} \circledast_m \bar\Phi^{k'}_{\underline{n}_{k+2}}(\mathrm{tr}_y\, Z_{2n_{k+2}-m})\frac{e^{-\mathscr{H}_{2n_{k+2}-m}}}{(2\pi)^{\frac{d(2n_{k+2}-m)}{2}}}\,\mathrm{d}Z_{2,2n_{k+2}-m}\,\mathrm{d}v_1 \leq \frac{\mu^{m-1}}{n_{k+2}^m}\left(\frac{\|h\|_0}{(\mu\mathfrak{d})^{n_{k+2}-1}}C^{n_{k+2}}\right)^2\delta^2\varepsilon^{\mathfrak{a}}\theta^{(n_{k+2}-n_k-2)_+}t^{n_k-1+n_{k+2}}. \tag{8.3}$$

Using the estimations (8.2) and (8.3), one obtains

$$\begin{aligned}\left|\mathbb{E}_\varepsilon\left[\mu^{-\frac12}\sum_{\underline{i}_{n_{k+2}}}\Phi^{k'}_{\underline{n}_{k+2}}(\mathbf{Z}_{\underline{i}_{n_{k+2}}}(t_s))\,\zeta^0_\varepsilon(g)\mathbb{1}_{\Upsilon_\varepsilon}\right]\right)\right| &\leq \|g\|_0\|h\|_0C^{n_{k+2}}\Big((\tfrac{\delta}{\mathfrak{d}})^2(\tfrac{\theta}{\mathfrak{d}})^{(n_{k+2}-n_k-2)_+}(\tfrac{t}{\mathfrak{d}})^{n_k-1}\varepsilon^{\frac{2\mathfrak{a}+1}{2}} \\ &\quad+\Big((\tfrac{\delta}{\mathfrak{d}})^2(\tfrac{\theta}{\mathfrak{d}})^{(n_{k+2}-n_k-2)_+}(\tfrac{t}{\mathfrak{d}})^{n_{k+2}+n_k-1}\varepsilon^{\mathfrak{a}}\Big)^{\frac12}\Big) \\ &\leq \delta\varepsilon^{\frac{\mathfrak{a}}{2}}\|h\|_0\|g\|_0C^{n_{k+2}}(\tfrac{t^2\theta}{\mathfrak{d}^3})^{\frac{(n_{k+2}-n_k-2)_+}{2}}(\tfrac{t}{\mathfrak{d}})^{n_k},\end{aligned}$$

as $\varepsilon^{\frac12}/\mathfrak{d} \to 0$.

Using that $\frac{Ct\theta}{\mathfrak{d}^2} \leq 1$ and $K'\delta = \theta \leq 1$, we can sum on k, k' and $\underline{n}_{k+2}$

$$\begin{aligned}\left|G_\varepsilon^{\mathrm{rec},2}(t)\right| &\leq \sum_{\substack{1\leq k\leq K-1\\1\leq k'\leq K'}}\sum_{\substack{n_1\leq\cdots\leq n_k\\ n_j-n_{j-1}\leq 2^j}}\sum_{n_{k+2}\geq n_{k+1}\geq n_k}\|g\|_0\|h\|_0\delta\varepsilon^{\frac{\mathfrak{a}}{2}}(\tfrac{Ct}{\mathfrak{d}})^{n_k}(\tfrac{Ct\theta}{\mathfrak{d}^2})^{\frac{(n_{k+2}-n_k-2)_+}{2}} \\ &\leq \|g\|_0\|h\|_0K'\delta\varepsilon^{\frac{\mathfrak{a}}{2}}K^{K^2}(\tfrac{Ct}{\mathfrak{d}})^{2^K} \\ &\leq \|g\|_0\|h\|_0\varepsilon^{\frac{\mathfrak{a}}{2}}(\tfrac{C't}{2\mathfrak{d}})^{2^K}\end{aligned}$$

This concludes the proof of (8.1).

Proof of (8.2). We recall that

$$\bar\Phi^{k'}_{\underline{n}_{k+2}}(Z_{n_{k+2}}) := \frac{1}{(n_{k+2})!}\sum_{\sigma\in\mathfrak{S}_{n_{k+2}}} \Phi^{>,\delta}_{n_{k+1},n_{k+2}}\Psi^{,t-t_s-\delta}_{\underline{n}_{k+1}}[h]\left(Z_{\sigma[n_{k+2}]}\right)$$

In $\Psi^{0,t-t_s-\delta}_{\underline{n}_{k+1}}\Phi^{0,\delta}_{n_{k+1},n_{k+2}}[h]\left(Z_{[n_{k+2}]}\right)$ we see three sets of indices:

- 1 the last particle,
- $[1,n_{k+1}]$ the set of particles in "final" tree pseudotrajectory development,
- $[n_{k+1}+1,n_{k+2}]$ the particles added in the first time interval.

Any permutation σ that sends $[1, n_{k+1}]$ and $[n_{k+1}+1, n_{k+2}]$ onto themselves stabilizes the function $\Psi^{0,t-t_s-\delta}_{\underline{n}_{k+1}} \Phi^{0,\delta}_{n_{k+1},n_{k+2}}[h]$. Hence, $\bar{\Phi}^{k'}_{\underline{n}_{k+2},p}(Z_p)$ is equal to

$$\frac{(n_{k+1}-1)!(n_{k+2}-n_{k+1})!}{(n_{k+2})!} \sum_{\substack{\omega_1 \sqcup \omega_2 = [n_{k+2}] \\ |\omega_1| = n_{k+1} \\ q_1 \in \omega_1}} \Phi^{>,\delta}_{n_{k+1},n_{k+2}} \Psi^{,t-t_s-\delta}_{\underline{n}_{k+1}}[h]\left(z_{q_1}, Z_{\omega_1 \setminus \{q_1\}}, Z_{\omega_2}\right).$$

Let us develop $\Psi^{0,t-t_s-\delta}_{\underline{n}_{k+1}} \Phi^{0,\delta}_{n_{k+1},n_{k+2}}[h]$. For $(s_i)_i \in \{\pm 1\}^{n_{k+1}-1}$, (ω_1, ω_2) a partition of $[n_{k+2}]$ and $(\lambda_1, \cdots, \lambda_{\mathbf{l}})$ a partition of $[n_{k+2}]$ with $\omega_1 \subset \lambda_1$, we define the pseudotrajectory $\bar{\mathsf{Z}}(\tau, Z_{n_{k+2}}, q_1, \omega_1, \omega_2, (s_i)_i, (\lambda_j)_j)$ by

- for $\tau \leq \delta$,
$$\bar{\mathsf{Z}}(\tau) := \mathsf{Z}(\tau, Z_{\omega_1}, (\lambda_j)_j)$$
- for $\tau > \delta$, the particle of ω_3 are removed and
$$\bar{\mathsf{Z}}_{\omega_1}(\tau) := \mathsf{Z}(\tau - \delta, \bar{\mathsf{Z}}_{\omega_1}(\delta), \{q_1\}, (s_i)_i).$$

Then $\bar{\Phi}^{k'}_{\underline{n}_{k+2}}(Z_{n_{k+2}})$ is equal to

$$\frac{1}{(n_{k+2})!} \sum_{\substack{\omega_1 \sqcup \omega_2 = [n_{k+2}] \\ |\omega_1| = n_{k+1} \\ q_1 \in \omega_1}} \sum_{(s_i)_i} \left(\sum_{\mathbf{l}=1}^{n} \sum_{\lambda_1 \supset \omega_1} \sum_{\substack{(\lambda_2, \cdots, \lambda_{\mathbf{l}}) \\ \in \mathscr{P}^{\mathbf{l}-1}_{\omega_2 \setminus \lambda_1}}} \right) h\left(\bar{\mathsf{z}}_{\omega_1}(k\theta + k'\delta)\right)$$
$$\times \mathbb{1}_{\substack{\bar{\mathsf{Z}}(\cdot) \text{ has a} \\ \text{pathology} \\ \text{on } [0,\delta]}} \left(\prod_{i=1}^{n_{k+2}-1} s_i \; \mathbb{1}_{\mathscr{R}^{0,k\theta+(k-1)\delta}_{\omega_1,(s_i)_i}(\bar{\mathsf{Z}}_{\omega_2}(\delta))} \prod_{i=1}^{k} \mathbb{1}_{\mathfrak{n}(t-i\theta)=n_i} \right)$$
$$\times \left(\mathscr{O}_l(Z_{\lambda_1}, \cdots, Z_{\lambda_{\mathbf{l}}}) \mathbb{\Delta}^{[m]}_{|\lambda_1|}(Z_{\lambda_1}) \prod_{i=2}^{\mathbf{l}} \mathbb{\Delta}_{|\lambda_i|}(Z_{\lambda_i}) \right).$$

The functions $\mathscr{O}$ and $\mathbb{\Delta}$ are defined in Definitions 3.1 and 3.2 (in the definition, they are defined on a time interval $[0,t]$; here, they are defined on $[0,\delta]$).

The function $\mathscr{O}_l(Z_{\lambda_1}, \cdots, Z_{\lambda_{\mathbf{l}}})$ can be bounded by the Penrose's tree inequality (see for example [BGSRS20, Jan]),

$$|\mathscr{O}_l(Z_{\lambda_1}, \cdots, Z_{\lambda_{\mathbf{l}}})| = \left| \sum_{C \in \mathscr{C}([\mathbf{l}])} \prod_{(i,j) \in E(C)} -\mathbb{1}_{\lambda_i \mathring{\sim} \lambda_j} \right| \leq \sum_{T \in \mathscr{T}([\mathbf{l}])} \prod_{(i,j) \in E(T)} \mathbb{1}_{\lambda_i \mathring{\sim} \lambda_j}.$$

The set $\mathscr{T}([\mathbf{l}])$ is the set of minimally connected graphs with vertices $[\mathbf{l}]$.

We explain now how to take advantage of the pathology of $\bar{\mathsf{Z}}(\cdot)$.

Definition 8.1. *For $r \geq 3$, we define the set $\mathscr{O}_r$ as*

$$\mathscr{O}_r := \Big\{ Z_r \in \mathbb{D}^r, \; \exists (\varpi_1, \cdots, \varpi_{\mathbf{l}}), \text{ the collision graph of } \mathsf{Z}_r(\cdot, Z_z, (\varpi_1, \cdots, \varpi_{\mathbf{l}})) \text{ on } [0,\delta] \text{ is connected, and the pseudotrajectory has a pathology} \Big\}. \tag{8.4}$$

We recall that a pathology can be an overlap, a multiple encounter , or a recollision (see Definition 3.7).

For $r = 2$, we define

$$\mathscr{O}_2 := \{|x_1 - x_2| \leq \varepsilon\} \cup \{|(x_1 - x_2) + \delta(v_1 - v_2)| \leq \varepsilon\}. \tag{8.5}$$

Finally, for $\varpi \subset [n_{k+2}]$, the set $\mathscr{O}_\varpi$ is defined as

$$\mathscr{O}_\varpi := \left\{ Z_{n_{k+2}} \in \mathbb{D}^{n_{k+2}}, \; Z_\varpi \in \mathscr{O}_{|\varpi|} \right\}. \tag{8.6}$$

The $\mathscr{O}_\varpi$ allows to control the recollision condition

$$\mathbb{1}_{\substack{\bar{\mathsf{Z}}(\cdot) \text{ has a} \\ \text{pathology} \\ \text{on } [0,\delta]}} \leq \sum_{\varpi \subset [n_{k+2}]} \mathbb{1}_{\mathscr{O}_\varpi}.$$

This leads to the following bound on $\left|\bar{\Phi}^{k'}_{\underline{n}_{k+2}}(Z_{n_{k+2}})\right|$:

$$(8.7)\quad \frac{\|h\|_0}{(n_{k+2})!}\sum_{\substack{\omega_1\sqcup\omega_2=[n_{k+2}]\\ |\omega_1|=n_{k+1}\\ q_1\in\omega_1}}\sum_{\substack{\varpi\subset[n_{k+2}]\\ |\varpi|\geq 2}}\sum_{(s_i)_i}\left(\sum_{\mathbf{l}=1}^{n}\sum_{\lambda_1\supset\omega_1}\sum_{\substack{(\lambda_2,\cdots,\lambda_{\mathbf{l}})\\ \in\mathscr{P}^{\mathbf{l}-1}_{\omega_2\setminus\lambda_1}}}\mathbb{1}_{\mathscr{R}^{0,k\theta+(k-1)\delta}_{\{q_1\},(s_i)_i}}(\bar{\mathsf{Z}}_{\omega_1}(\delta))\mathbb{1}_{\mathfrak{n}(k'\delta)=n_k}\right.$$
$$\left.\times\,\mathbb{1}_{\mathscr{O}_\varpi}\sum_{T\in\mathscr{T}([\mathbf{l}])}\prod_{(i,j)\in E(T)}\mathbb{1}_{\lambda_i\mathring{\sim}\lambda_j}\mathbb{M}^{[n_{k+1}]}_{|\lambda_1|}(Z_{\lambda_1})\prod_{i=2}^{\mathbf{l}}\mathbb{M}_{|\lambda_i|}(Z_{\lambda_i})\right).$$

Note that the right hand-side is invariant under translation. Thus, one can fix $x_1=0$ and integrate with respect to the other variables.

We introduce a partition to control the pseudo-trajectory in the time interval $[0,\delta]$.

Definition 8.2 (Possible cluster partition)**.** *Given $Z_{n_{k+2}}\in\mathbb{D}^{n_{k+2}}$, we construct the graph G with vertices $[n_{k+2}]$. The pair (i,j) is an edge of G if and only if there exists $\tilde{\omega}\subset[n_{k+2}]$ and $(\tilde{\lambda}_1,\cdots,\tilde{\lambda}_\ell)$ a partition of $\tilde{\omega}$ such that the collision graph of $\mathrm{Z}(\cdot,Z_{\tilde{\omega}},\tilde{\lambda}_1,\cdots,\tilde{\lambda}_\ell)$ on time interval $[0,\delta]$ is connected. We introduce $\underline{\rho}:=(\rho_1,\cdots,\rho_{\mathbf{r}})$ the* possible cluster partition *as the set of the connected components of G.*

We define $\mathscr{D}^{\underline{\rho}}_{\varepsilon}\subset\mathbb{D}^{n_{k+2}}$ as the set such that $\underline{\rho}$ is the possible cluster partition. The $\left(\mathscr{D}^{\underline{\rho}}_{\varepsilon}\right)_{\underline{\rho}}$ form a partition of $\mathbb{D}^{n_{k+2}}$.

By definition of the potential cluster, a particle cannot interact with a particle of an other cluster for any time in $[0,\delta]$. Thus the systems ρ^i are isolated on $[0,\delta]$ and all the dynamics in $[0,\delta]$ are encoded inside the (ρ_i).

The parametrization of the pseudotrajectories is changed to a more adapted one. There exists a ρ_i containing ϖ. With a little loss of symmetry, one can suppose that it is ρ_1. In the same way, for any λ_j with $j\neq 1$ there exists some ρ_i containing λ_j. For any ρ_i

- $\underline{\omega}^i:=(\omega^i_1,\omega^i_2)$ the partition of ρ_i defined by $\omega^i_j:=\omega_j\cap\rho_i$, note that the set ω^i_1 cannot be empty,
- $\underline{\lambda}^i:=\{\lambda^i_1:=\lambda_1\cap\rho_i\}\cup\{\lambda_j \text{ for } j\geq 2 \text{ with } \lambda_j\subset\rho_i\}$ a partition of ρ_i,
- for $i\geq 1$, $\mathfrak{p}_i:=(\underline{\omega}^i,\underline{\lambda}^i)$,
- $\mathfrak{p}_1:=(\underline{\omega}^1,\underline{\lambda}^1,\varpi)$.

The set of possible $\mathfrak{p}_i$ is denoted $\mathfrak{P}(\rho_i)$. Because ρ_i is of size at most γ, there exists a constant C_γ depending only on γ such that $|\mathfrak{P}(\rho_i)|\leq C_\gamma$. For a fix partition $\underline{\rho}$, the map $(\underline{\omega},\varpi,\underline{\lambda})\mapsto(\mathfrak{p}_i)_i$ is onto.

The possible cluster partition also contains the overlap: if we denote two dynamical clusters λ_j and $\lambda_{j'}$ with $j,j'\geq 2$, there exists a ρ_i containing both, and if $\lambda_j\subset\rho_i$ has an overlap with λ_1, then λ_j has an overlap with λ^i_1. This last property allows us to rewrite the overlap cumulant: for any $Z_{n_{k+2}}$ in $\mathscr{D}^{\underline{\rho}}_{\varepsilon}$,

$$\left|\psi_{\mathbf{l}}(Z_{\lambda_1},\cdots,Z_{\lambda_{\mathbf{l}}})\right|\leq\sum_{T\in\mathscr{T}([\mathbf{l}])}\prod_{(i,j)\in E(T)}\mathbb{1}_{\lambda_i\mathring{\sim}\lambda_j}\leq\prod_{i=1}^{\mathbf{r}}\sum_{T_i\in\mathscr{T}([|\underline{\lambda}^i|])}\prod_{(j,j')\in E(T_i)}\mathbb{1}_{\lambda^i_j\mathring{\sim}\lambda^i_{j'}}\leq\prod_{i=1}^{\mathbf{l}}\left|\mathscr{T}([|\underline{\lambda}^i|])\right|.$$

The right-hand side is bounded using that

$$\left|\mathscr{T}([|\underline{\lambda}^i|])\right|\leq|\underline{\lambda}^i|^{|\underline{\lambda}^i|-2}\leq|\rho_i|^{|\rho_i|}$$

(see section 2 of [BGSRS20]). As the symmetric conditioning imposes that $|\rho_i|\leq\gamma$, the cumulants $\left|\psi_{\mathbf{l}}(Z_{\lambda_1},\cdots,Z_{\lambda_{\mathbf{l}}})\right|$ are smaller than $(\gamma^\gamma)^{n_{k+2}}$.

We now have the following bound

$$(8.8)\qquad \left|\bar{\Phi}^{k'}_{\underline{n}_{k+2},p}(Z_p)\right|\leq\frac{\gamma^{n_{k+2}}\|h\|_0}{(n_{k+2})!}\sum_{q_1=1}^{n_{k+2}}\sum_{\mathbf{r}=1}^{n_{k+2}}\sum_{\underline{\rho}\in\mathscr{P}^{\mathbf{r}}_{n_{k+2}}}\sum_{\substack{(s_i)_i\\ \underline{\mathfrak{p}}\in\prod_i\mathfrak{P}(\rho_i)}}\mathbb{1}_{\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{q_1,(s_i)_i}}(Z_{n_{k+2}})\prod_{i=1}^{\mathbf{r}}\Delta_{\mathfrak{p}_i}(Z_{\rho_i}),$$

where we denote

$$\Delta_{\mathfrak{p}_1}(Z_{\rho_1}):=\mathbb{1}_{\mathscr{O}_\varpi}\mathbb{1}_{Z_{\rho_1}\text{ form a possible cluster}},$$
$$\forall i\geq 2,\ \Delta_{\mathfrak{p}_i}(Z_{\rho_i}):=\mathbb{1}_{Z_{\rho_i}\text{ form a possible cluster}}\quad\text{and}$$
$$\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{q_1,(s_i)_i}:=\left\{Z_p\in\mathscr{D}^{\underline{\rho}}_{\varepsilon},\,\bar{\mathsf{Z}}_{\omega_1}(\delta)\in\mathscr{R}^{0,k\theta+(k-1)\delta}_{\{q_1\},(s_i)_i}\right\}.$$

The same method as in [BGSRS23a] is used to control the condition $\mathbb{1}_{\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{((s_i)_i}}$.

For a pseudotrajectory $\bar{Z}_{n_{k+2}}(\tau)$, consider its collision graph $\mathscr{G}^{[0,t-t_s]}_{\omega_1\cup\omega_2}$. Then, we can construct the graph G by identifying in $\mathscr{G}^{[0,t-t_s]}_{\omega_1\cup\omega_2}$ the particles in the same cluster ρ_i. Finally we can construct the *clustering trees* $T^> := (\nu_i, \bar{\nu}_i)_{1\le i\le \mathbf{r}-1}$ where the i-th clustering collision in G happens between cluster ρ_{ν_i} and $\rho_{\bar{\nu}_i}$.

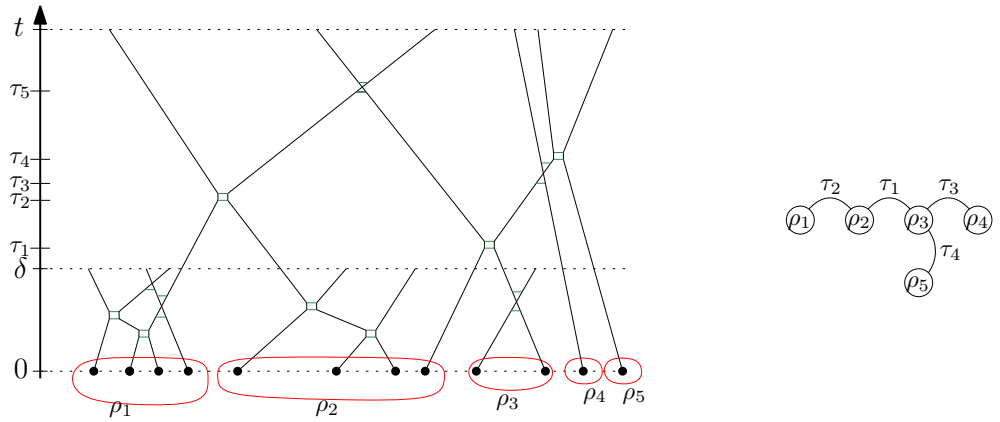

FIGURE 10. Example of construction of the clustering stets.

We need to count the number of clustering collisions of $T^>$ happening between time δ and time $k'\delta$. If $\mathbf{r} > n_k$, all the $\mathbf{r}-1$ collisions in $T^>$ cannot correspond to the $n_k - 1$ collisions of the time interval $[k'\delta, \theta]$. Thus, at least $(\mathbf{r}-n_k)_+$ collisions happen during $[\delta, k'\delta] \subset [0, 2\theta]$.

One needs a different representation of collision graphs. Let L_0 be equal to $\{\{1\}, \cdots, \{\mathbf{r}\}\}$. The L_i and $(\nu_{(i)}, \bar{\nu}_{(i)})$ are constructed sequentially: suppose that $L_{i-1} = (c_1, \cdots, c_l)$, the (c_j) forming a partition of $[1, r]$. The i-th collision happens between cluster $\nu_i \in c_a$ and $\bar{\nu}_i \in c_b$. Then:

- $L_i := \big(L_{i-1} \setminus \{c_a, c_b\}\big) \cup \{c_a \cup c_b\}$,
- $\{\nu_{(i)}, \bar{\nu}_{(i)}\} := \{c_a, c_b\}$ with $\max \nu_{(i)} < \max \bar{\nu}_{(i)}$.

The $(\nu_{(i)}, \bar{\nu}_{(i)})_i$ define a partition of $\mathscr{T}^>([\mathbf{r}])$ (the set of ordered trees on $[1, \mathbf{r}]$).

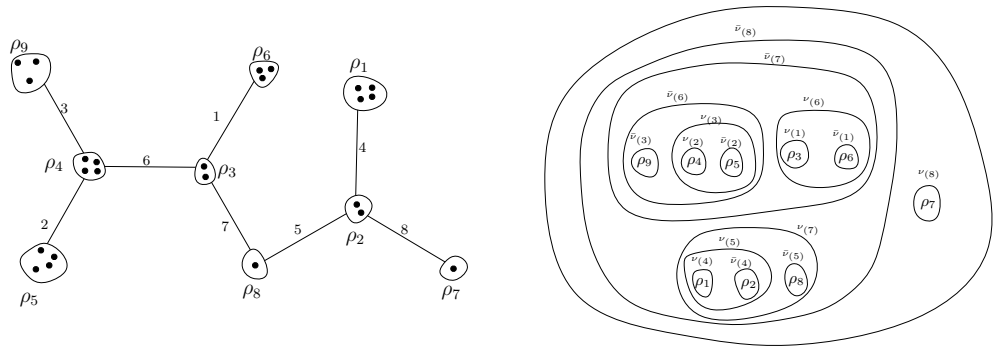

FIGURE 11. An example of construction of the representation $(\nu_{(i)}, \bar{\nu}_{(i)})_i$ from a clustering graph. The graph

We performed the following change of variables:

$$\forall i \in \{1, \cdots, \mathbf{r}-1\},\ \hat{x}_i := x_{\min \nu_{(i)}} - x_{\min \bar{\nu}_{(i)}},\ \tilde{X}_i := \mathrm{tr}_{-x_{\min \rho_i}} X_{\rho_i},$$

$$X_{2,n_{k+2}} \mapsto (\hat{x}_1 \cdots, \hat{x}_{\mathbf{r}-1}, \tilde{X}_1, \cdots, \tilde{X}_{\mathbf{r}}).$$

The condition $\mathscr{R}^{\rho,\mathfrak{p}}_{q_1,(s_i)_i}$ is integrated first with respect to $(\hat{x}_1, \cdots, \hat{x}_{\mathbf{r}-1})$, where relative positions inside a cluster $\tilde{X}_i$ are kept constant. The $(\Delta_{\mathfrak{p}_i})_i$ will be summed later with respect to the $(\tilde{X}_i)_i$.

Fix τ_{i+1} the time of the $(i+1)$-th clustering collision, and the relative positions $\hat{x}_{i-1}, \cdots, \hat{x}_1$. We define the i-th clustering set

$$B_i := \bigcup_{\substack{q\in\omega_{2,\nu_{(i)}}\\ \bar{q}\in\omega_{1,\bar{\nu}_{(i)}}}} B_i^{q,\bar{q}}$$

$$\text{with} \qquad \omega_{1,\nu_{(i)}} := \bigcup_{j\in\nu_{(i)}} \omega_1^j, \qquad \omega_{1,\bar{\nu}_{(i)}} := \bigcup_{\bar{\jmath}\in\bar{\nu}_{(i)}} \omega_2^{\bar{\jmath}},$$

$$B_i^{q,\bar{q}} := \Big\{ \hat{x}_i \;\Big|\; \exists \tau_i \in [0, \tau_{i+1}\wedge T_i],\ |\bar{\mathsf{x}}_{\bar{q}}(\tau_i) - \bar{\mathsf{x}}_{\bar{q}}(\tau_i)| = \varepsilon \Big\}$$

and $T_i := 2\theta$ for the $(\mathbf{r}-n_k)_+$ first collisions, t else. We used that $\omega_1 \cup \omega_2$ is the set of particles existing after time δ.

Up to time τ_i the curves $\bar{\mathsf{x}}_q$ and $\bar{\mathsf{x}}_{\bar{q}}$ are independent. Hence, we can perform the change of variables $\hat{x}_i \mapsto (\tau_i, \eta_i)$, where τ_i is the minimal collision time and

$$\eta_i := \frac{\bar{\mathsf{x}}_{\bar{q}}(\tau_i) - \bar{\mathsf{x}}_q(\tau_i)}{|\bar{\mathsf{x}}_{\bar{q}}(\tau_i) - \bar{\mathsf{x}}_q(\tau_i)|}.$$

The Jacobian of this diffeomorphism is $\varepsilon^{d-1}|(\bar{\mathsf{v}}_{\bar{q}}(\tau_i) - \bar{\mathsf{v}}_q(\tau_i))\cdot\eta_i|\,\mathrm{d}\tau_i\,\mathrm{d}\eta_i$.

As the particles in $\omega_{1,\nu_{(i)}}$ and $\omega_{1,\bar{\nu}_{(i)}}$ are isolated during $[\delta,\tau_i]$, their energies are conserved. The sum of relative velocities can be bounded by

$$\begin{aligned}\sum_{\substack{q\in\omega_{1,\nu_{(i)}}\\ \bar{q}\in\omega_{1,\bar{\nu}_{(i)}}}} |\bar{\mathsf{v}}_{\bar{q}}(\tau_i) - \bar{\mathsf{v}}_q(\tau_i)| &\le |\bar{\mathsf{V}}_{\omega_{1,\nu_{(i)}}}(\tau_i)|\,|\omega_{1,\nu_{(i)}}|^{1/2}|\omega_{1,\bar{\nu}_{(i)}}| + |\bar{\mathsf{V}}_{\omega_{1,\bar{\nu}_{(i)}}}(\tau_i)|\,|\omega_{1,\bar{\nu}_{(i)}}|^{1/2}|\omega_{1,\nu_{(i)}}|\\ &\le \Big(|\omega_{1,\nu_{(i)}}| + |\bar{\mathsf{V}}_{\omega_{1,\nu_{(i)}}}(\tau_i)|^2\Big)\Big(|\omega_{1,\bar{\nu}_{(i)}}| + |\bar{\mathsf{V}}_{\omega_{1,\bar{\nu}_{(i)}}}(\tau_i)|^2\Big).\end{aligned}$$

Using the same method as in the proof of 5.5,

$$\frac{1}{2}\left|\bar{\mathsf{V}}_{\omega_{1,\nu_{(i)}}}(\tau_i)\right|^2 \le \mathscr{H}_{|\omega_{1,\nu_{(i)}}|}\left(\bar{\mathsf{Z}}_{\omega_{1,\nu_{(i)}}}(\delta)\right) \le \mathscr{H}_{|\lambda_{1,\bar{\nu}_{(i)}}|}\left(\bar{\mathsf{Z}}_{\lambda_{1,\bar{\nu}_{(i)}}}(\delta)\right),$$

where we denote

$$\lambda_{1,\bar{\nu}_{(i)}} := \bigcup_{\bar{\jmath}\in\bar{\nu}_{(i)}} \lambda_1^j.$$

At time δ, the particles in two different clusters cannot interact (by definition of a possible cluster),

$$\mathscr{H}_{|\lambda_{1,\nu_{(i)}}|}(\bar{\mathsf{Z}}_{\lambda_{1,\nu_{(i)}}}(\delta)) = \sum_{j\in\nu_{(i)}} \mathscr{H}_{|\lambda_1^j|}(\bar{\mathsf{Z}}_{\lambda_1^j}(\delta)) = \sum_{j\in\nu_{(i)}} \mathscr{H}_{|\lambda_1^j|}(\bar{\mathsf{Z}}_{\lambda_1^j}(0)) \le \sum_{j\in\nu_{(i)}} \mathscr{H}_{|\rho_j|}(Z_{\rho_j}).$$

We conclude that

$$\sum_{\substack{q\in\omega_{1,\nu_{(i)}}\\ \bar{q}\in\omega_{1,\bar{\nu}_{(i)}}}} |\bar{\mathsf{v}}_{\bar{q}}(\tau_i) - \bar{\mathsf{v}}_q(\tau_i)| \le 4 \sum_{\substack{\nu_i\in\nu_{(i)}\\ \bar{\nu}_i\in\bar{\nu}_{(i)}}} \Big(|\rho_{\nu_i}| + \mathscr{H}_{|\rho_{\nu_i}|}(Z_{\rho_{\nu_i}})\Big)\Big(|\rho_{\bar{\nu}_i}| + \mathscr{H}_{|\rho_{\bar{\nu}_i}|}(Z_{\rho_{\bar{\nu}_i}})\Big). \tag{8.9}$$

This gives the following bound on $|B_i|$ (using the Boltzmann-Grad scaling $\mu\mathfrak{d}\varepsilon^{d-1}=1$)

$$\begin{aligned}|B_i| &\le \frac{C}{\mu\mathfrak{d}}\int_0^{t_{i+1}\wedge T_i} \mathrm{d}\tau_i \sum_{q,\bar{q}} |\bar{\mathsf{v}}_q(\tau_i) - \bar{\mathsf{v}}_{\bar{q}}(\tau_i)|\\ &\le \frac{C}{\mu\mathfrak{d}} \sum_{\substack{\nu_i\in\nu_{(i)}\\ \bar{\nu}_i\in\bar{\nu}_{(i)}}} \Big(|\rho_{\nu_i}| + \mathscr{H}_{|\rho_{\nu_i}|}(Z_{\rho_{\nu_i}})\Big)\Big(|\rho_{\bar{\nu}_i}| + \mathscr{H}_{|\rho_{\bar{\nu}_i}|}(Z_{\rho_{\bar{\nu}_i}})\Big)\int_0^{t_{i+1}\wedge T_i} \mathrm{d}\tau_i.\end{aligned}$$

Permuting the product and the sum,

$$\begin{aligned}\sum_{(\nu_{(i)},\bar{\nu}_{(i)})_i} \prod_{i=1}^{\mathbf{r}-1} \sum_{\substack{\nu_i\in\nu_{(i)}\\ \bar{\nu}_i\in\bar{\nu}_{(i)}}} &\Big(|\rho_{\nu_i}| + \mathscr{H}_{|\rho_{\nu_i}|}(Z_{\rho_{\nu_i}})\Big)\Big(|\rho_{\bar{\nu}_i}| + \mathscr{H}_{|\rho_{\bar{\nu}_i}|}(Z_{\rho_{\bar{\nu}_i}})\Big)\\ &= \sum_{(\nu_i,\bar{\nu}_i)_i} \prod_{i=1}^{\mathbf{r}-1} \Big(|\rho_{\nu_i}| + \mathscr{H}_{|\rho_{\nu_i}|}(Z_{\rho_{\nu_i}})\Big)\Big(|\rho_{\bar{\nu}_i}| + \mathscr{H}_{|\rho_{\bar{\nu}_i}|}(Z_{\rho_{\bar{\nu}_i}})\Big).\end{aligned}$$

Using that

$$\forall a,b\in\mathbb{N},\ \frac{(a+b)!}{a!b!} \le 2^{a+b},$$

we have

$$\int_0^t \mathrm{d}t_{\mathbf{r}-1}\cdots\int_0^{t_2\wedge T_2}\mathrm{d}t_1 \le \frac{t^{n_k\wedge\mathbf{r}-1}}{(n_k\wedge\mathbf{r}-1)!}\frac{\theta^{(\mathbf{r}-n_k)_+}}{((\mathbf{r}-n_k)_+)!} \le 2^{n_{k+2}}\frac{t^{n_k\wedge\mathbf{r}-1}\theta^{(\mathbf{r}-n_k)_+}}{(\mathbf{r}-1)!}.$$

We can now sum up on the clustering collisions:

$$\begin{aligned}\int \mathbb{1}_{\mathscr{R}^{\rho,\mathfrak{p}}_{(s_i)_i}}\mathrm{d}\hat{x}_1\cdots\mathrm{d}\hat{x}_{\mathbf{r}-1} &\le \sum_{(\nu_{(i)},\bar{\nu}_{(i)})}\int\mathrm{d}\hat{x}_1'\mathbb{1}_{B_1}\int\mathrm{d}\hat{x}_2'\mathbb{1}_{B_2}\cdots\int\mathrm{d}\hat{x}_{\mathbf{r}-1}\mathbb{1}_{B_{\mathbf{r}-1}}\\ &\le\left(\frac{C}{\mu\mathfrak{d}}\right)^{\mathbf{r}-1}\int_0^t\mathrm{d}t_{\mathbf{r}-1}\cdots\int_0^{t_2\wedge T_2}\mathrm{d}t_1\sum_{(\nu_i,\bar{\nu}_i)_i}\prod_{i=1}^{\mathbf{r}-1}\Big(|\rho_{\nu_i}|+\mathscr{H}_{|\rho_{\nu_i}|}(Z_{\rho_{\nu_i}})\Big)\Big(|\rho_{\bar{\nu}_i}|+\mathscr{H}_{|\rho_{\bar{\nu}_i}|}(Z_{\rho_{\bar{\nu}_i}})\Big)\\ &\le\left(\frac{2C}{\mu\mathfrak{d}}\right)^{\mathbf{r}-1}\frac{t^{n_k\wedge\mathbf{r}-1}\theta^{(\mathbf{r}-n_k)_+}}{(\mathbf{r}-1)!}\sum_{(\nu_i,\bar{\nu}_i)_i}\prod_{i=1}^{\mathbf{r}-1}\Big(|\rho_{\nu_i}|+\mathscr{H}_{|\rho_{\nu_i}|}(Z_{\rho_{\nu_i}})\Big)\Big(|\rho_{\bar{\nu}_i}|+\mathscr{H}_{|\rho_{\bar{\nu}_i}|}(Z_{\rho_{\bar{\nu}_i}})\Big).\end{aligned}$$

We can forget the order of the edges of $T^{>}=(\nu_i,\bar{\nu}_i)_i$, which gives a factor $\mathbf{r}!$. Secondly, denoting $d_i(G)$ the degree of the vertex i in a graph G and $\mathscr{T}([\mathbf{r}])$ the set of minimally (not oriented) connected graphs on $[1,\mathbf{r}]$, we can write the preceding inequality as

$$\int\mathbb{1}_{\mathscr{R}^{\rho,\mathfrak{p}}_{q_1,(s_i)_i}}\mathrm{d}\hat{x}_1\cdots\mathrm{d}\hat{x}_{\mathbf{r}-1}\le\left(\frac{2C}{\mu\mathfrak{d}}\right)^{\mathbf{r}-1}t^{n_k\wedge\mathbf{r}-1}\theta^{(\mathbf{r}-n_k)_+}\sum_{T\in\mathscr{T}([\mathbf{r}])}\prod_{i=1}^{r}\Big(|\rho_i|+\mathscr{H}_{|\rho_i|}(Z_{\rho_i})\Big)^{d_i(T)}.$$

For $A,B>0$, $x\in\mathbb{R}^+$, there exists a constant $C>0$ such that

$$(A+x)^B e^{-\frac{x}{4}}\le\left(\tfrac{4B}{e}\right)^B e^{\frac{A}{4}}.$$

We use this inequality and that on $\mathscr{D}^{\rho}_{\rho}$, $\mathscr{H}_n(Z_n)=\sum_i\mathscr{H}_{|\rho_i|}(Z_{\rho_i})$ to bound

$$\begin{aligned}&\int\mathbb{1}_{\mathscr{R}^{\rho,\mathfrak{p}}_{(s_i)_i}}e^{-\frac{1}{2}\mathscr{H}_{n_{k+2}}(Z_{n_{k+2}})}\mathrm{d}\hat{x}_1\cdots\mathrm{d}\hat{x}_{\mathbf{r}-1}\\ &\qquad\le\left(\frac{C}{\mu\mathfrak{d}}\right)^{\mathbf{r}-1}t^{n_k\wedge r-1}\theta^{(\mathbf{r}-n_k)_+}\sum_{T\in\mathscr{T}([\mathbf{r}])}\prod_{i=1}^{\mathbf{r}}\Big(|\rho_i|+\mathscr{H}_{|\rho_i|}(Z_{\rho_i})\Big)^{d_i(T)}e^{-\frac{1}{2}\sum_{i=1}^{\mathbf{r}}\mathscr{H}_{|\rho_i|}(Z_{\rho_i})}\\ &\qquad\le\tilde{C}^{n_{k+2}}\frac{t^{n_k\wedge\mathbf{r}-1}\theta^{(\mathbf{r}-n_k)_+}}{(\mu\mathfrak{d})^{\mathbf{r}-1}}\sum_{T\in\mathscr{T}([\mathbf{r}])}\prod_{i=1}^{\mathbf{r}}d_i(T)^{d_i(T)}\,.\end{aligned}$$

As the sum of the $d_i(T)$ is equal to $2\mathbf{r}-2$, we have by convexity of $x\mapsto x\log x$

$$\sum_{i=1}^{\mathbf{r}}d_i(T)\log d_i(T)\le\mathbf{r}\frac{\sum_{i=1}^{\mathbf{r}}d_i(T)}{\mathbf{r}}\log\frac{\sum_{i=1}^{\mathbf{r}}d_i(T)}{\mathbf{r}}\le(2\mathbf{r}-2)\log 2$$

and $|\mathscr{T}([\mathbf{r}])|$ is equal to $\mathbf{r}^{\mathbf{r}-2}$,

$$\begin{aligned}&\int\mathbb{1}_{\mathscr{R}^{\rho,\mathfrak{p}}_{(s_i)_i}}e^{-\frac{1}{2}\mathscr{H}_{n_{k+2}}(Z_{n_{k+2}})}\mathrm{d}\hat{x}_1\cdots\mathrm{d}\hat{x}_{\mathbf{r}-1}\\ &\qquad\le\left(\frac{C}{\mu\mathfrak{d}}\right)^{\mathbf{r}-1}t^{n_k\wedge r-1}\theta^{(\mathbf{r}-n_k)_+}\sum_{T\in\mathscr{T}([\mathbf{r}])}\prod_{i=1}^{\mathbf{r}}\Big(|\rho_i|+\mathscr{H}_{|\rho_i|}(Z_{\rho_i})\Big)^{d_i(T)}e^{-\frac{1}{2}\sum_{i=1}^{\mathbf{r}}\mathscr{H}_{|\rho_i|}(Z_{\rho_i})}\\ &\qquad\le\tilde{C}'^{n_{k+2}}\frac{t^{n_k\wedge\mathbf{r}-1}\theta^{(\mathbf{r}-n_k)_+}}{(\mu\mathfrak{d})^{\mathbf{r}-1}}(\mathbf{r}-1)!\,.\end{aligned}$$

We can now integrate the condition $\Delta_{\mathfrak{p}_i}(Z_{\rho_i})$. The particles in Z_{ρ_i} have to form a possible cluster. Because clusters are of size at most γ,

$$\int_{\mathbb{T}^{|\rho_1|-1}\times(\mathbb{R}^d)^{|\rho_1|}}\Delta_{\mathfrak{p}_1}(Z_{\rho_1})\frac{e^{-\frac{1}{2}\mathscr{H}_{|\rho_1|}(Z_{\rho_1})}}{(2\pi)^{d|\rho_1|/2}}\,\mathrm{d}\tilde{X}_i\,\mathrm{d}V_{\rho_1}\le\frac{C_r\delta^{\max\{2,|\rho_1|-1\}}}{(\mu\mathfrak{d})^{|\rho_1|-1}}\varepsilon^{\mathfrak{a}},\tag{8.10}$$

$$\int_{\mathbb{T}^{|\rho_i|-1}\times(\mathbb{R}^d)^{|\rho_i|}}\Delta_{\mathfrak{p}_i}(Z_{\rho_i})\frac{e^{-\frac{1}{2}\mathscr{H}_{|\rho_i|}(Z_{\rho_i})}}{(2\pi)^{d|\rho_i|/2}}\,\mathrm{d}\tilde{X}_i\,\mathrm{d}V_{\rho_i}\le C_\gamma\left(\frac{\delta}{\mathfrak{d}\mu}\right)^{|\rho_i|-1},\tag{8.11}$$

The second inequality is a clustering estimation, similar to the ones threatened in the proof of (5.2). In the first inequality, we use recollision estimates as in the proof of (7.2). The proofs are given in Appendix B.4.

Integrating the $\Delta_{\mathfrak{p}_i}$ leads to

$$\int \mathbb{1}_{\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{(s_i)_i}}(Z_{n_{k+2}}) \prod_{i=1}^{\mathbf{r}} \Delta_{\mathfrak{p}_i}(Z_{\rho_i}) \frac{e^{-\mathscr{H}_{n_{k+2}}}}{(2\pi)^{dn_{k+2}/2}} \,\mathrm{d}X_{2,n_{k+2}} \,\mathrm{d}V_{n_{k+2}}$$
$$\leq (\mathbf{r}-1)! C^{n_{k+2}} \frac{t^{n_k \wedge \mathbf{r}-1} \theta^{(\mathbf{r}-n_k)_+}}{(\mu\mathfrak{d})^{\mathbf{r}-1}} \prod_{i=1}^{\mathbf{r}} \left(\frac{\delta}{\mathfrak{d}\mu}\right)^{|\rho_i|-1} \frac{\delta^{\max 2,|\rho_1|-1}}{(\mu\mathfrak{d})^{|\rho_1|-1}} \varepsilon^{\mathfrak{a}}.$$

Any particle removed at time δ has a clustering collision during $[0,\delta]$. Therefore, $\sum_{i=1}^{\mathbf{r}}(|\rho_i|-1)$ is bigger than $n_{k+2}-n_k$. In addition, we have chosen θ bigger than δ so

$$\int \mathbb{1}_{\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{(s_i)_i}}(Z_{n_{k+2}}) \prod_{i=1}^{\mathbf{r}} \Delta_{\mathfrak{p}_i}(Z_{\rho_i}) \frac{e^{-\mathscr{H}_{n_{k+2}}(Z_{n_{k+2}})}}{(2\pi)^{\frac{dn_{k+2}}{2}}} \,\mathrm{d}X_{2,n_{k+2}} \,\mathrm{d}V_{n_{k+2}}$$
$$\leq (\mathbf{r}-1)! \frac{C^{n_{k+2}}}{(\mu\mathfrak{d})^{n_{k+2}-1}} t^{n_k-1} \theta^{(n_{k+2}-n_k-2)_+} \delta^2 \varepsilon^{\mathfrak{a}}.$$

We sum now on the parameters $(s_i)_i$ and $(\mathfrak{p}_i)$. Because the sizes of the possible clusters are bounded by γ, the $|\mathfrak{P}(\rho_i)|$ are smaller than some $C_\gamma > 0$ depending only on γ. The number of collision parameters $(s_i)_i$ is equal to $2^{n_{k+2}}$ and

$$\int \left|\bar{\Phi}^{k'}_{\underline{n}_{k+2}}(Z_{n_{k+2}})\right| M^{\otimes n_{k+2}} \,\mathrm{d}X_{2,n_{k+2}} \,\mathrm{d}V_{n_{k+2}}$$
$$\leq \frac{\|h\|(CC_\gamma 4\gamma)^{n_{k+2}}}{n_{k+2}!(\mu\mathfrak{d})^{n_{k+2}-1}} t^{n_k-1}\theta^{(n_{k+2}-n_k-2)_+}\delta^2\varepsilon^{\mathfrak{a}} \sum_{r=1}^{n_{k+2}} \sum_{\underline{\rho}\in\mathscr{P}^r_{n_{k+2}}} (\mathbf{r}-1)!\,.$$

The last sums can be bounded by

$$\frac{1}{n_{k+2}!}\sum_{r=1}^{n_{k+2}} \sum_{\underline{\rho}\in\mathscr{P}^r_{n_{k+2}}} (\mathbf{r}-1)! = \frac{1}{n_{k+2}!}\sum_{\mathbf{r}=1}^{n_{k+2}} \sum_{\substack{k_1+\cdots+k_{\mathbf{r}}=n_{k+2}\\ k_i\geq 1}} \frac{n_{k+2}!}{k_1!\cdots k_{\mathbf{r}}!}\frac{(\mathbf{r}-1)!}{\mathbf{r}!}$$
$$\leq \sum_{\mathbf{r}=1}^{n_{k+2}} \sum_{\substack{k_1+\cdots+k_{\mathbf{r}}=n_{k+2}\\ k_i\geq 1}} \frac{1}{k_1!\cdots k_{\mathbf{r}}!} \leq e^{n_{k+2}}$$

This ends the proof of the first inequality. □

Proof of (8.3). As the $\bar{\Phi}^{k'}_{\underline{n}_{k+2}}$ are symmetric, it is sufficient to study

$$\left|\bar{\Phi}^{k'}_{\underline{n}_{k+2}}(Z_{[n_{k+2}]})\bar{\Phi}^{k'}_{\underline{n}_{k+2}}(Z_{[n_{k+2}+1-m,2n_{k+2}-m]})\right|. \tag{8.12}$$

The bound (8.7) leads to

$$(8.12) \leq \frac{\|h\|^2}{(n_{k+2}!)^2} \sum_{\substack{(q_1,\omega_1\omega_2)\\(q_1'\omega_1'\omega_2')}} \sum_{\varpi\subset[n_{k+2}]} \sum_{\substack{(s_i)_i\\(s_i')_i}} \sum_{\substack{(\lambda_1,\cdots,\lambda_{\mathbf{l}})\\(\lambda_1',\cdots,\lambda_{\mathbf{l}}')}} \mathbb{1}_{\mathscr{R}^{0,t_s-\delta}_{\{q_1\},(s_i)_i}}(\bar{\mathsf{Z}}_{\omega_1}(\delta))\mathbb{1}_{\mathscr{R}^{0,t_s-\delta}_{\{q_1\},(s_i')_i}}(\bar{\mathsf{Z}}'_{\omega_2'}(\delta))$$
$$\times\left(\sum_{T\in\mathscr{T}([\mathbf{l}])}\prod_{(i,j)\in E(T)}\mathbb{1}_{\lambda_i\mathring{\sim}\lambda_j}\mathbb{M}^{[n_{k+1}]}_{|\lambda_1|}(Z_{\lambda_1})\prod_{i=2}^{\mathbf{l}}\mathbb{M}_{|\lambda_i|}(Z_{\lambda_i})\right)$$
$$\times\left(\sum_{T'\in\mathscr{T}([\mathbf{l}'])}\prod_{(i,j)\in E(T')}\mathbb{1}_{\lambda'_i\mathring{\sim}\lambda'_j}\mathbb{M}^{[n_{k+1}]}_{|\lambda_1'|}(Z_{\lambda_1'})\prod_{i=2}^{\mathbf{l}}\mathbb{M}_{|\lambda_i'|}(Z_{\lambda_i'})\right)\mathbb{1}_{\mathbf{n}(k'\delta)=n_k}\mathbb{1}_{\mathscr{O}_\varpi}.$$

where we have denoted $\mathscr{T}(\sigma)$ the set of connected and simply connected graphs with vertices σ, a finite set. The sets $\mathscr{O}_\varpi$ have been defined in Definition 8.1. In addition, we sum on all parameters such that

- $q_1 \in [n_{k+2}]$ and $q_1' \in [n_{k+2}+1-m, 2n_{k+2}-m]$,
- $\omega_1 \sqcup \omega_2 = [n_{k+2}]$, $\omega_1' \sqcup \omega_2' = [n_{k+2}+1-m, 2n_{k+2}-m]$, $q_1 \in \omega_1$, $q_1 \in \omega_1$ and $|\omega_1| = |\omega_1'| = n_{k+1}$,
- $\lambda_1 \supset \omega_1$, $\lambda_1' \supset \omega_1'$, $(\lambda_2,\cdots,\lambda_{\mathbf{l}})$ an unordered partition of $[n_{k+2}]\setminus\omega_1$ and $(\lambda_2',\cdots,\lambda_{\mathbf{l}}')$ an unordered partition of $[n_{k+2}+1-m, 2n_{k+2}-m]\setminus\omega_1'$.

The pseudotrajectory $\bar{Z}(\tau)$ (respectively $\bar{Z}'(\tau)$) begins with coordinates $Z_{[n_{k+2}]}$ (respectively with coordinates $Z_{[n_{k+2}+1-m,2n_{k+2}-m]}$) with parameters $(q_1,\omega_1,\omega_2,(\lambda_1,\cdots,\lambda_{\mathbf{l}}))$ (respectively $(q'_1,\omega'_1,\omega'_2,(\lambda'_1,\cdots,\lambda'_{\mathbf{l}'}))$) in the same way than in the proof of (8.2).

Note that the right hand-side of (8.12) is invariant under translation. Thus, one can fix $x_1=0$ and integrate with respect to the other variables.

For a position $Z_{2n_{k+2}-m}$, we consider $\underline{\rho}:=(\rho_1,\cdots,\rho_{\mathbf{r}})$ the possible clusters. As in the previous section, with a little loss of symmetry, one can suppose that $\varpi_1\subset\rho_1$. We can then construct the parameters $\mathfrak{p}_1:=(\underline{\omega}^1,\underline{\omega}'^1,\underline{\lambda}^1,\underline{\lambda}'^1,\varpi)$, $(\mathfrak{p}_i)_{i\geq 2}:=\big((\underline{\omega}^i,\underline{\omega}'^i,\underline{\lambda}^i,\underline{\lambda}'^i)\big)_{i\geq 2}$:

- $\underline{\omega}^i:=(\omega_1^i,\omega_2^i)$ is a partition of $\rho_i\cap[n_{k+2}]$ defined by $\omega_j^i:=\omega_j\cap\rho_i$,
- $\underline{\omega}'^i:=({\omega'}_1^i,{\omega'}_3^i)$ is a partition of $\rho_i\cap[n_{k+2}+1-m,2n_{k+2}+m]$ defined by ${\omega'}_j^i:=\omega'_j\cap\rho_i$,
- $\underline{\lambda}^i:=\{\lambda_1^i:=\lambda_1\cap\rho_i\}\cup\{\lambda_j \text{ for } j\geq 2 \text{ with } \lambda_j\subset\rho_i\}$ a partition of $[n_{k+2}]\cap\rho_i$,
- $\underline{\lambda'}^i:=\{{\lambda'}_1^i:=\lambda'_1\cap\rho_i\}\cup\{\lambda'_j \text{ for } j\geq 2 \text{ with } \lambda'_j\subset\rho_i\}$ a partition of $[n_{k+2}+1-m,2n_{k+2}+m]\cap\rho_i$.

We denote now $\mathfrak{P}(\rho_i)$ the new set of possible parameters $\mathfrak{p}_i$ (this will not create a conflict of notations with the previous section). Because each cluster ρ_i is of size at most γ, $|\mathfrak{P}(\rho_i)|$ is bounded by some constant C_γ depending only on γ. Defining

$$\Delta_{\mathfrak{p}_1}(Z_{\rho_1}):=\mathbb{1}_{\substack{Z_{\rho_1}\text{ form a}\\ \text{possible cluster}}}\mathbb{1}_{\mathscr{O}_\varpi},\quad \forall i\geq 2,\ \Delta_{\mathfrak{p}_i}(Z_{\rho_i}):=\mathbb{1}_{\substack{Z_{\rho_i}\text{ form a}\\ \text{possible cluster}}}$$

$$\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{(s_i)_i,(s'_i)_i}:=\Big\{Z_{2n_{k+2}-m}\in\mathscr{D}^{\underline{\rho}}_{\bar{\varepsilon}},\ \bar{\mathsf{Z}}_{\omega_1\cup\omega_2\cup\omega_3}(\delta)\in\mathscr{R}^{0,k\theta+(k-1)\delta}_{\omega_1,(s_i)_i},\bar{\mathsf{Z}}'_{n_{k+2}}(\delta)\in\mathscr{R}^{0,k\theta+(k-1)\delta}_{\omega'_1,(s'_i)_i}\Big\},$$

we have as in the inequality (8.8)

$$(8.12)\leq\frac{\gamma^{2(\gamma-2)n_{k+2}}\|h\|_0^2}{(n_{k+2}!)^2}\sum_{\mathbf{r}=1}^{2n_{k+2}-m}\sum_{\underline{\rho}\in\mathscr{P}^r_p}\sum_{\substack{(s_i)_i,(s'_i)_i\\ \underline{\mathfrak{p}}\in\prod_i\mathfrak{P}(\rho_i)}}\mathbb{1}_{\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{(s_i)_i,(s'_i)_i}}(Z_{2n_{k+2}-m})\prod_{i=1}^{\mathbf{r}}\Delta_{\mathfrak{p}_i}(Z_{\rho_i}).$$

Note that, for at least one i, $\underline{\varpi}^i$ is not empty. We are now constructing a clustering tree in order to estimate $\mathscr{R}^{\underline{\rho},\underline{\mathfrak{p}}}_{(s_i)_i,(s'_i)_i}$.

Consider the collision graph associated with the first pseudotrajectory $\mathscr{G}^{[0,t-t_s]}_{\omega_1\cup\omega_2}$ and the graph associated with the second one $\mathscr{G}^{[0,t-t_s]}_{\omega'_1\cup\omega'_2}$. Merge them and identify vertices in the same cluster ρ_i. Keeping only the first clustering collisions, we obtain the oriented tree $T^>:=(\nu_i,\bar{\nu}_i)_{1\leq i\leq\mathbf{r}-1}$. Note that these clustering collisions can happen in the first or second pseudotrajectories.

As in the proof of (8.2) we have to bound the number of collisions of $T^>$ in the time interval $[0,2\tau]$. There are at most $(n_k-1+n_{k+2}-m)$ collisions during $[(k'+1)\delta,t-t_s]$ (n_k-1 for the first pseudotrajectory, and we have to connect $n_{k+2}-m$ particles in the second). Thus, there are at least $(\mathbf{r}-(n_k-1+n_{k+2}-m))_+$ clustering collisions in $[\delta,(k'+1)\delta]\subset[0,2\tau]$.

We explain quickly how to estimate the i-th collision. As in the previous paragraph, we construct the modified tree parameters $(\nu_{(i)},\bar{\nu}_{(i)})$ and the change of variables

$$\forall i\in\{1,\cdots,\mathbf{r}-1\},\ \hat{x}_i:=x_{\min\nu_{(i)}}-x_{\min\bar{\nu}_{(i)}},\ \tilde{X}_i:=\mathrm{tr}_{x_{\min\rho_i}}X_{\rho_i},$$

$$X_{2,2n_{k+2}-m}\mapsto(\hat{x}_1\cdots,\hat{x}_{\mathbf{r}-1},\tilde{X}_1,\cdots,\tilde{X}_r),$$

and we integrate on the $(\hat{x}_i)$.

The clustering set B_i is defined as follows: fix t_{i+1} the time of the $(i+1)$-th clustering collision, and the relative positions $\hat{x}_{i-1},\cdots,\hat{x}_1$. We define the i-th clustering set

$$B_i:=\bigcup_{\substack{q\in\bigcup_{j\in\nu_{(i)}}\rho_j\\ \bar{q}\in\bigcup_{\bar{\jmath}\in\bar{\nu}_{(i)}}\rho_{\bar{\jmath}}}}\Big(B_i^{q,\bar{q}}\cup B_i^{\prime q,\bar{q}}\Big)$$

with

$$B_i^{q,\bar{q}}:=\Big\{\hat{x}_i\ \Big|\ \exists\tau_i\in[0,t_{i+1}\wedge T_i],\ |\bar{\mathsf{x}}_{\bar{q}}(\tau_i)-\bar{\mathsf{x}}_q(\tau_i)|=\varepsilon\Big\},$$

where $T_i := 2\theta$ for the $(\mathbf{r} - n_k)_+$ first collisions, and t else. The set $B_i'^{q,q'}$ is defined in the same way for the other pseudotrajectory. We can apply the estimates of the previous paragraph:

$$\int \mathbb{1}_{B_i} \,\mathrm{d}\hat{x}_i \leq \frac{2C}{\mu\mathfrak{d}} \sum_{\substack{\nu_i \in \nu_{(i)} \\ \bar{\nu}_i \in \bar{\nu}_{(i)}}} \Big(|\rho_{\nu_i}| + \mathscr{H}_{|\rho_{\nu_i}|}(Z_{\rho_{\nu_i}})\Big)\Big(|\rho_{\bar{\nu}_i}| + \mathscr{H}_{|\rho_{\bar{\nu}_i}|}(Z_{\rho_{\bar{\nu}_i}})\Big) \int_0^{\tau_{i+1}\wedge T_i} \mathrm{d}\tau_i.$$

In this way, we end up with the same situation as in the estimate of (8.2), and we can apply the same strategy:

$$\begin{aligned}\int |\bar{\Phi}^{k'}_{\underline{n}_{k+2}}(Z_{k+2})\bar{\Phi}^{k'}_{\underline{n}_{k+2}}(Z_{n_{k+2}+1-m,2n_{k+2}-m})| \frac{e^{-\mathscr{H}_{2_{k+2}-m}(Z_{2n_{k+2}-m})}}{(2\pi)^{2n_{k+2}-m}} \,\mathrm{d}X_{2,2n_{k+2}-m} \,\mathrm{d}V_{2n_{k+2}-m} \\ \leq \frac{(2n_{k+2}-m)!\|h\|^2}{(n_{k+2}!)^2(\mu\mathfrak{d})^{2n_{k+2}-m-1}} C^{n_{k+2}}\delta^2\varepsilon^{\mathfrak{a}}\tau^{(n_{k+2}-n_k-2)_+}t^{n_k-1+n_{k+2}-m} \\ \leq \frac{\mu^{m-1}}{n_{k+2}^m}\left(\frac{\|h\|}{(\mu\mathfrak{d})^{n_{k+2}-1}}\tilde{C}^{n_{k+2}}\right)^2 \delta^2\varepsilon^{\mathfrak{a}}\theta^{(n_{k+2}-n_k-2)_+}t^{n_k-1+n_{k+2}}\end{aligned}$$

which concludes the proof. □

Appendix A. The linearized Boltzmann operator without cut-off

In this section, we fix the dimension $d = 3$. We construct the linearized Boltzmann operator associated with the power law $1/r^s$, $s > 1$ and we explain where the scaling $\mathfrak{d}_{s,\alpha} = \alpha^{2/s}$ comes from.

We begin with a change of variables in the definition of the Boltzmann operator $\mathscr{L}_\alpha$. For (v, v_*, ν), we define

$$\vec{\rho} := \tfrac{\nu\wedge(v-v_*)}{|v-v_*|} \in \mathrm{span}(v - v_*)^\perp \tag{A.1}$$

the *impact parameters*, with the Jacobian

$$((v - v_*)\cdot\nu)_+ \,\mathrm{d}\nu \to |v - v_*| \,\mathrm{d}\vec{\rho}.$$

This allows us to redefine the post-collisional velocities (v', v_*) for an interaction potential $\mathscr{U}$

$$\begin{cases} (v', v_*') := \lim\limits_{t\to\infty}(v_a(t), v_b(t)) \\ \dfrac{\mathrm{d}}{\mathrm{d}t}(x_a, x_b) = (v_a, v_b), \ \dfrac{\mathrm{d}}{\mathrm{d}t}(v_a, v_b) = \alpha(-\nabla\mathscr{U}(x_b - x_a), \nabla\mathscr{U}(x_b - x_a)) \\ \lim\limits_{t\to-\infty}(v_a(t), v_b(t)) =: (v, v_*), \ (v_a - v_b)\times(x_a - x_b) = |v - v_*|\vec{\rho}. \end{cases} \tag{A.2}$$

With this definition, the scattering map can easily be defined for a not compactly supported decreasing potential.

For $\lambda > 0$, we make the change of coordinate

$$(t, x_a, x_b, v_a, v_b) \mapsto (\tilde{t}, \tilde{x}_a, \tilde{x}_b, v_a, v_b) := (\lambda t, \lambda x_a, \lambda x_b, v_a, v_b)$$

In the new coordinates, the equations of motion become

$$\tfrac{\mathrm{d}}{\mathrm{d}\tilde{t}}(\tilde{x}_a, \tilde{x}_b) = (\tilde{v}_a, \tilde{v}_b), \ \tfrac{\mathrm{d}}{\mathrm{d}\tilde{t}}(\tilde{v}_a, \tilde{v}_b) = \alpha\left(-\nabla\mathscr{U}\left(\tfrac{\tilde{x}_a - \tilde{x}_b}{\lambda}\right), \nabla\mathscr{U}\left(\tfrac{\tilde{x}_a - \tilde{x}_b}{\lambda}\right)\right)$$

Hence, the post-collisional parameters associated with $(v, v_*, \vec{\rho})$ and potential $\mathscr{U}$ are the same as the ones associated with $(v, v_*, \lambda\vec{\rho})$ and potential $\mathscr{U}(\cdot/\lambda)$.

Performing the change of variable $\vec{\rho} \to \alpha^{-1/s}\vec{\rho}$ in the collisional operator gives

$$\alpha^{2/s}\mathscr{L}_{\alpha\mathscr{V}}g = \mathscr{L}_{\mathscr{U}_\alpha}g, \text{ where } \mathscr{U}_\alpha(r) := \alpha\mathscr{V}(r\alpha^{1/s}) = \frac{f\left(r\alpha^{\frac{1}{s}}\right)}{r^s}.$$

This new potential converges when $\alpha \to 0$ to $\mathscr{U}^s(r) := 1/r^s$. It is natural to guess the convergence of the operators

$$\frac{1}{\alpha^{2/s}}\mathscr{L}_\alpha \to \mathscr{L}_{\mathscr{U}^s}.$$

Section 3 of [LBW22] is dedicated to the proof of this claim.

Appendix B. Geometrical estimates

B.1. Estimation the length scattering time.

Lemma B.1. *Let $\mathscr{V}$ an interaction potential that is radial and supported in a ball of diameter ε. We consider two particles, 1 and 2, with initial coordinates*

$$(x_a(0), v_a(0)) = (0, v_1),\ (x_b(0), v_b(0)) = (\varepsilon\nu, v_2),\ \nu \in \mathbb{S}^{d-1},\ (v_1 - v_2)\cdot\nu < 0,$$

following the Hamiltonian dynamics linked to

$$H := \tfrac{|v_a|^2+|v_b|^2}{2} + \mathscr{V}(\tfrac{x_a - x_b}{\varepsilon}).$$

Then, the scattering time is bound by

$$[\tau] := \inf\{\tau > 0\ |x_a(\tau) - x_b(\tau)| > \varepsilon\} \lesssim \frac{\varepsilon|v_1 - v_2|}{|\nu \times (v_1 - v_2)|^2}. \tag{B.1}$$

Proof. The motion equation are written as

$$\begin{aligned} \tfrac{d}{dt}(x_a + x_b) &= (v_a + v_b) \quad \tfrac{\mathrm{d}}{\mathrm{d}t}(v_a + v_b) = 0 \\ \tfrac{\mathrm{d}}{\mathrm{d}t}(x_a - x_b) &= (v_a - v_b) \quad \tfrac{\mathrm{d}}{\mathrm{d}t}(v_a - v_b) = -\tfrac{2}{\varepsilon}\nabla\mathscr{V}\left(\tfrac{x_a - x_b}{\varepsilon}\right). \end{aligned}$$

Hence, $[\tau]$ does not depend on $v_1 + v_2$.

We use the impact parameter $\rho := \frac{|\varepsilon\nu\times(v_1-v_2)|}{|v_1-v_2|}$ defined in (A.1). The time $[\tau]$ is given by (see chapter 8 of [GSRT13])

$$[\tau] = \frac{2}{|v_1 - v_2|}\int_{r_{\min}}^{\varepsilon/2} \frac{\mathrm{d}r}{\sqrt{1 - \frac{\rho^2}{r^2} - 2\frac{\mathscr{V}(r/\varepsilon)}{|v_1-v_2|^2}}}, \tag{B.2}$$

with $r_{\min}$ defined by

$$1 - \tfrac{\rho^2}{r_{\min}^2} - 2\frac{\mathscr{V}(r_{\min}/\varepsilon)}{|v_1 - v_2|^2} = 0.$$

We begin by performing the change of variables

$$u^2 := \frac{\rho^2}{r^2} + 2\frac{\mathscr{V}(r/\varepsilon)}{|v_1 - v_2|^2} \tag{B.3}$$

which implies

$$[\tau] = \frac{2}{|v_1 - v_2|}\int_{\rho/\varepsilon}^{1} \frac{u\,\mathrm{d}u}{\sqrt{1-u^2}} \frac{r}{\frac{\rho^2}{r^2} - \frac{r\mathscr{V}'(r/\varepsilon)}{\varepsilon|v_1-v_2|^2}} \tag{B.4}$$

Using that $\mathscr{V}'$ is non positive and that $u \geq \frac{\rho}{r}$, $r \leq \varepsilon$

$$[\tau] \leq \frac{2}{|v_1 - v_2|}\int_{\rho/\varepsilon}^{1} \frac{\mathrm{d}u}{\sqrt{1-u^2}} \frac{u\varepsilon^3}{\rho^2} \lesssim \frac{\varepsilon^3}{|v_1 - v_2|\rho^2} \lesssim \frac{\varepsilon|v_1 - v_2|}{|\nu \times (v_1 - v_2)|^2}.$$

□

B.2. Proof of Proposition 7.2.

The goal of this section is proving the following estimations (6.8) and (7.4):

Lemma B.2. *Let $\mathscr{V}$ be an interaction potential that is radial, decreasing, and supported in a ball of radius ε.*

There exists a $C > 0$ independent of $\mathscr{V}$ such that for $t = n_k\theta$

$$\begin{aligned} &\sum_{T=(q_i,\bar{q}_i,s_i)_{i\leq n_k-1}} \int_{\mathfrak{T}_{n_k}\times\mathfrak{G}_T^0} \left(1 - \mathbb{1}_{\mathfrak{G}_T^0}\right) \Lambda_T(V_{n_K}, \nu_{[n_k-1]})\,\mathrm{d}\nu_{[n_k-1]}\,\mathrm{d}\tau_i M^{\otimes n_k} dV_{n_k} \\ &\quad + \sum_{\bar{T}} \int_{\mathfrak{T}_{\underline{n}_{k+2}}\times\mathfrak{G}^{>,t-t_s,\bar{T}}_{\{q\},\omega,(s_i)_i}} \prod_{i=1}^{n_{k+2}-1} |(\mathrm{v}^{\varepsilon}_{q_i}(\tau_i) - \mathrm{v}^{\varepsilon}_{q_i'}(\tau_i))\cdot\nu_i|\,\mathrm{d}\nu_i\,\mathrm{d}\tau_i M^{\otimes n_{k+2}} dV_{n_{k+2}} \\ &\qquad\qquad \leq C^{n_K}(n_k)^{n_k}\theta^{(n_k - n_{k-1} - 2)_+} t^{n_k}\varepsilon^{1/5}. \end{aligned} \tag{B.5}$$

The two estimations can be performed in the same way, and we will only treat the first one. The proof of this lemma is an adaptation of the proof of Lemma 8 of [PSS14]. The estimation is not optimal. For example, the factor $\varepsilon^{1/4}$ can be replaced in the hard spheres setting by $\varepsilon|\log\varepsilon|^r$ for some constant $r>1$ (see for example [BGSRS23b]). However, optimal estimates use the upper bound of the collision kernel of hard spheres. Such bounds are verified for more general potential and certainly not in the limit $\alpha\to 0$. The proof of [PSS14] (which is adapted from it) is more robust.

Proof. We need to avoid

- an *overlap*: there exists a time $\tau\in(0,t)\cap\delta\mathbb{Z}$ and two particles q and q' such that
$$|\mathsf{x}_q(\tau)-\mathsf{x}_{q'}(\tau)|\le\varepsilon,$$
- a *recollision*: there exists a time $\tau\in[0,t]$ and two particles q and q' such that $\tau\notin\{\tau_1,\cdots,\tau_{n_k-1}\}$ and
$$|\mathsf{x}_q(\tau)-\mathsf{x}_{q'}(\tau)|=\varepsilon \text{ and } (\mathsf{x}_q(\tau)-\mathsf{x}_{q'}(\tau))\cdot(\mathsf{v}_q(\tau)-\mathsf{v}_{q'}(\tau))<0.$$

We begin with the estimation of overlap, which is easier. As the i-th collision between particles $(q_i,\bar{q_i})$ can last only a time $\frac{\varepsilon\left|\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|}{\left|\left(\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right)\times\nu_i\right|^2}$ (see (B.1)), there can be an overlap only if there is some $\tau\in\delta\mathbb{Z}\cap[0,t]$ such that τ_i is in the interval

$$I_\varepsilon(\tau,V_{n_K},\nu_{[n_k-1]}):=\left[\tau-\frac{\varepsilon\left|\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|}{\left|\left(\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right)\times\nu_i\right|^2},\tau\right]$$

Hence, the set of parameters leading to an overlap is smaller than

$$\sum_{\substack{\tau\in\delta\mathbb{Z}\cap[0,t]\\1\le i\le n_k-1}}\sum_T\int_{\mathfrak{T}_{n_k}\times\mathfrak{G}_T^0}\mathbb{1}_{I_\varepsilon(\tau,V_{n_K},\nu_{[n_k-1]})}(\tau_i)\Lambda_T(V_{n_K},\nu_{[n_k-1]})\,\mathrm{d}\nu_{[n_k-1]}\,\mathrm{d}\tau_{[n_k-1]}M^{\otimes n_k}dV_{n_k}$$
$$\le\frac{C^{n_k}t^{1+n_k-1}\theta^{(n_k-n_{k-1}-1)_+}}{\delta n_k^{n_k}}\sum_{1\le i\le n_k-1}\sum_T\int_{\mathfrak{G}_T^0}t\wedge\frac{\varepsilon\left|\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|}{\left|\left(\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right)\times\nu_i\right|^2}\Lambda_T(V_{n_K},\nu_{[n_k-1]})\,\mathrm{d}\nu_{[n_k-1]}M^{\otimes n_k}dV_{n_k}$$

We can apply the same estimates as in Lemma 6.1 and for the terms (6.11). We conclude that the set of overlap has a measure smaller than

$$\frac{\varepsilon^{1/2}C^{n_k}t^{n_{k-1}+1}\theta^{(n_k-n_{k-1}-1)_+}}{\delta n_k^{n_k}}.$$

We treat the recollision now. We denote $\mathscr{G}$ the collision graph, and we define $\mathscr{G}^{[0,\tau)}$ the subgraph of $\mathscr{G}$ with edges

$$\{(q,\bar{q})_{\tau',\sigma}\in E(G),\ \tau'<\tau\}.$$

If the first recollision involves particles q and q' at time τ_{rec}, we consider $\omega\subset[n_{k+2}]$ the connected components of $\{q,q'\}$ in the collision graph $\mathscr{G}^{[0,\tau_{\mathrm{rec}})}$. Before the time τ_{rec}, the pseudotrajectory $\mathsf{Z}_\omega^\varepsilon(\tau)$ and its formal limit $\mathsf{Z}_\omega(\tau)$ are close up to a translation, and using Lemma 6.3, there exists a $y_0\in\mathbb{T}$ such that

$$\forall\tau\in[0,\tau_{\mathrm{rec}}],\ |\mathsf{X}_\omega^0(\tau)-\mathrm{tr}_{y_0}\mathsf{X}_\omega^\varepsilon(\tau)|\le 2n_K\mathbb{V}\sum_{i=1}^{n_{k+2}-1}1\wedge\frac{\varepsilon\left|\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|}{\left|\left(\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right)\times\nu_i\right|^2}.$$

Hence, there is a recollision if at time $\tau_{\mathrm{rec}}\in[0,t]\setminus\bigcup_i\{\tau_i\}$,

$$\exists\, q,q'\in[n_k-1]\text{ such that }|\mathsf{x}_q^0(\tau_{\mathrm{rec}})-\mathsf{x}_{q'}^0(\tau_{\mathrm{rec}})|\le\varepsilon+\sum_{i=1}^{n_{k+2}-1}1\wedge\varepsilon\frac{2n_K\mathbb{V}\left|\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|}{\left|\left(\mathsf{v}_{q_i}(\tau_i^-)-\mathsf{v}_{\bar{q}_i}(\tau_i^-)\right)\times\nu_i\right|^2}.\tag{B.6}$$

We can study only the limiting flow and defining a contact as "there exists a time τ_{rec} such that (B.6) is verified": we have

$$\mathfrak{T}_{\underline{n}_k}\times\mathfrak{G}_T^0\setminus\mathfrak{T}_{\underline{n}_k}\times\mathfrak{G}_T^\varepsilon\subset\mathfrak{T}_{\underline{n}_k}\times\mathfrak{G}_T^0\cap\{\text{at least one contact}\}.$$

The first step is to forbid the collisions that last too long. We define $\mathscr{E}_1 \subset \mathfrak{T}_{\underline{n}_k} \times \mathfrak{G}_T^0$ as (for some $c_1, c_2 \in (0,1)$)

$$\mathscr{E}_1 := \Bigg\{ \forall i \leq n_k - 1, \left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{q_i}(\tau_i^-)\right| \min\{1, (\tau_i - \tau_{i-1}), (\tau_{i+1} - \tau_i)\} \geq \frac{\varepsilon^{c_1}}{n_k^2 \mathbb{V}} \\ \frac{\left|\left(\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right) \times \nu_i\right|^2}{\left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|} \min\{1, (\tau_i - \tau_{i-1}), (\tau_{i+1} - \tau_i)\} \geq \frac{\varepsilon^{c_2}}{n_k^2 \mathbb{V}} \Bigg\}.$$

In $\mathscr{E}_1$ there is a contact if there exists a time τ_{rec} such that

$$|\mathsf{x}_q^0(\tau_{\mathrm{rec}}) - \mathsf{x}_{q'}^0(\tau_{\mathrm{rec}})| \leq 3\varepsilon^{1-c_2}, \tag{B.7}$$

and

$$1 - \mathbb{1}_{\mathscr{E}_1} \leq \sum_{i=1}^{n_k - 1} \mathbb{1}_{\left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{q_i}(\tau_i^-)\right| \leq \frac{\varepsilon^{c_1}}{n_k^2 \mathbb{V}}} + \mathbb{1}_{\frac{\left|\left(\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right) \times \nu_i\right|^2}{\left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|} \leq \frac{\varepsilon^{c_2}}{n_k^2 \mathbb{V}}} \\ + \mathbb{1}_{|\tau_i - \tau_{i-1}| \leq \max\left(\frac{\varepsilon^{c_1}}{n_k^2 \mathbb{V} \left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{q_i}(\tau_i^-)\right|}, \frac{\varepsilon^{c_2} \left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|}{n_k^2 \mathbb{V} \left|\left(\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right) \times \nu_i\right|^2} \right)} \\ + \mathbb{1}_{|\tau_{i+1} - \tau_i| \leq \max\left(\frac{\varepsilon^{c_1}}{n_k^2 \mathbb{V} \left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{q_i}(\tau_i^-)\right|}, \frac{\varepsilon^{c_2} \left|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right|}{n_k^2 \mathbb{V} \left|\left(\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)\right) \times \nu_i\right|^2} \right)}.$$

Now, we place ourselves in $\mathscr{E}_1$ and we fix a collision tree T. The first contact happens at time τ_{rec} between particles q_{rec} and q'_{rec}. There exists a collision i_0, two disjoint sequences of collisions $(i_j)_{j \leq p}$ and $(i'_j)_{j \leq p'}$ and two sequences of particles $(a_j)_{j \leq p}$, $(a'_j)_{j \leq p'}$ such that

- $\forall j \geq 1$, $i_0 < i_j$ and $i_0 < i'_j$,
- $\forall j \geq 1$, $a_j \in \{q_{i_j}, q'_{i_j}\} \cap \{q_{i_{j-1}}, q'_{i_{j-1}}\}$ and $a'_j \in \{q_{i'_j}, q'_{i'_j}\} \cap \{q_{i'_{j-1}}, q'_{i'_{j-1}}\}$,
- if for $j < j'$, $a_j = a_{j'}$. Then for any $i \in [i_j, i_{j'}]$ such that $a_j \in \{q_i, q'_i\}$, we have $j \in \{i_j, i_{j+1}, \cdots, i_{j'}\}$, and similarly for the sequences $(i'_j)_{j \leq p'}$, $(a'_j)_{j \leq p'}$,
- $a_0 = q_{i_0}$, $a'_0 = q'_{i_0}$ and $\{a_p, a'_{p'}\} = \{q_{\mathrm{rec}}, q'_{\mathrm{rec}}\}$.

There is a contact if

$$\min_{\substack{s \in [-t,t]R \\ y_0 \in \mathbb{Z}^d}} \Bigg| y_0 + \sum_{j=1}^{p} \mathsf{v}_{a_j}(]\tau_{i_{j-1}}, \tau_{i_j}[)(\tau_{i_j} - \tau_{i_{j-1}}) + \mathsf{v}_{a_p}(]\tau_{i_p}, \tau_{\mathrm{rec}}[)(s - \tau_{i_p}) \\ - \sum_{j=1}^{p'} \mathsf{v}_{a'_j}(]\tau_{i'_{j-1}}, \tau_{i'_j}[)(\tau_{i'_j} - \tau_{i'_{j-1}}) + \mathsf{v}_{a'_{p'}}(]\tau_{i'_{p'}}, \tau_{\mathrm{rec}}[)(s - \tau_{i'_{p'}}) \Bigg| \leq \varepsilon^{1-c_2}. \tag{B.8}$$

In addition, if the $(\tau_i)_i$ and $(\tau'_i)_i$ verify $\tau_{i_0} < \tau_{i_1} < \cdots < \tau_{i_p} < \tau_{\mathrm{rec}}$ and $\tau_{i_0} < \tau_{i'_1} < \cdots < \tau_{i'_p} < \tau_{\mathrm{rec}}$, there is a contact if

$$\min_{\substack{s \in [\tau_{i_0+1}, t]R \\ y_0 \in \mathbb{Z}^d}} \Bigg| y_0 + \sum_{j=1}^{p} \mathsf{v}_{a_j}(]\tau_{i_{j-1}}, \tau_{i_j}[)(\tau_{i_j} - \tau_{i_{j-1}}) + \mathsf{v}_{a_p}(]\tau_{i_p}, \tau_{\mathrm{rec}}[)(s - \tau_{i_p}) \\ - \sum_{j=1}^{p'} \mathsf{v}_{a'_j}(]\tau_{i'_{j-1}}, \tau_{i'_j}[)(\tau_{i'_j} - \tau_{i'_{j-1}}) + \mathsf{v}_{a'_{p'}}(]\tau_{i'_{p'}}, \tau_{\mathrm{rec}}[)(s - \tau_{i'_{p'}}) \Bigg| \leq \varepsilon^{1-c_2}. \tag{B.9}$$

Remark B.1. *Note that we can perform the previous construction if $\tau_{\mathrm{rec}} > \tau_{i_0}$. If we study the overlap for particles of size ε, it is always the case. But for punctual particles, the pathology can happen before the first collision. Our proof can be adapted by taking i_0 such that $i_0 > i_j$ and $i_0 > i')_j$, making the change of variable $V_{n_k} \mapsto \mathsf{V}_{n_k}(\tau_{i_0}^-)$ and looking at everything backwards.*

Denoting

$$\Delta \mathsf{v}_j := \mathsf{v}_{a_{j+1}}(]\tau_{i_j}, \tau_{i_{j+1}}[) - \mathsf{v}_{a_j}(]\tau_{i_{j-1}}, \tau_{i_j}[), \; \Delta \mathsf{v}'_j := \mathsf{v}_{a'_{j+1}}(]\tau_{i'_j}, \tau_{i'_{j+1}}[) - \mathsf{v}_{a'_j}(]\tau_{i'_{j-1}}, \tau_{i'_j}[), \\ w_0 := \mathsf{v}_{a_0}(\tau_{i_0}^+) - \mathsf{v}_{a'_0}(\tau_{i_0}^+), \text{ and } w_{\mathrm{f}} = \mathsf{v}_{a_p}(]\tau_{i_p}, \tau_{\mathrm{rec}}[) - \mathsf{v}_{a'_{p'}}(]\tau_{i'_{p'}}, \tau_{\mathrm{rec}}[),$$

the equation (B.8) can be written as

$$\min_{\substack{s\in[-t,t]\\ y_0\in\mathbb{Z}^d}} \left| y_0 - \tau_{i_0} w_0 + \sum_{j=1}^{p} \Delta \mathsf{v}_j \tau_{i_j} - \sum_{j=1}^{p'} \Delta \mathsf{v}'_j \tau_{i'_j} + s w_{\mathrm{f}} \right| \leq \varepsilon^{1-c_2} \tag{B.10}$$

In the following, we denote $\hat{w}_{\mathrm{f}} := \frac{w_{\mathrm{f}}}{|w_{\mathrm{f}}|}$. We take the cross product with $\hat{w}_{\mathrm{f}}$ in (B.10):

$$\min_{y_0\in\mathbb{Z}^d} \left| (y_0 - \tau_{i_0} w_0) \times \hat{w}_{\mathrm{f}} + \sum_{j=1}^{p} \Delta \mathsf{v}_j \times \hat{w}_{\mathrm{f}} \tau_{i_j} - \sum_{j=1}^{p'} \Delta \mathsf{v}'_j \times \hat{w}_{\mathrm{f}} \tau_{i'_j} \right| \leq \varepsilon^{1-c_2}. \tag{B.11}$$

We fix y_0 such that the minimum is reached. We distinguish three cases:

(1) If there exists a $\ell_1 \in \mathbb{N}$ such that $|\Delta \mathsf{v}_{\ell_1} \times \hat{w}_{\mathrm{f}}| \geq \frac{\varepsilon^{c_3}}{n_k}$ (for some $c_3 \in (c_1, 1)$), $\tau_{i_{\ell_1}}$ has to be in an interval of length $\varepsilon^{1-c_2-c_3}$.

(2) Else, if y_0 is non zero, using that

$$w_0 = -\sum_{j=1}^{p} \Delta \mathsf{v}_j + \sum_{j=1}^{p'} \Delta \mathsf{v}'_j - w_{\mathrm{f}}$$

we have

$$|w_0 \times \hat{w}_{\mathrm{f}}| \leq \sum_{j=1}^{p} |\Delta \mathsf{v}_j \times \hat{w}_{\mathrm{f}}| + \sum_{j=1}^{p'} |\Delta \mathsf{v}_j \times \hat{w}_{\mathrm{f}}| \leq \varepsilon^{c_3}.$$

Hence $|y_0 \times w_f|$ has to be smaller than $2t\mathbb{V}\varepsilon^{c_3} + \varepsilon^{1-c_2}$. Combining with (B.11), we deduce that there exists a collision j, a vector $y_0 \in \mathbb{Z}^d$ with $|y_0| \leq n_K t\mathbb{V}$, and a couple of particles (q, q') such that $w_{\mathrm{f}} = \big(\mathsf{v}_q(\tau_j^-) - \mathsf{v}_{q'}(\tau_j^-)\big)$, and then

$$\big|\big(\mathsf{v}_q(\tau_j^-) - \mathsf{v}_{q'}(\tau_j^-)\big) \times y_0\big| \leq 2t\mathbb{V}\varepsilon^{c_3} + \varepsilon^{1-c_2}.$$

(3) We treat now the case $y_0 = 0$. We begin to show that there exists some ℓ' such that $|\Delta \mathsf{v}_{\ell'}|$ or $|\Delta \mathsf{v}'_{\ell'}|$ is bigger than $\frac{\varepsilon^{c_1}}{3n_k t}$ (in the following, we suppose that it is $|\Delta \mathsf{v}_{\ell'}|$).

Proof. We proceed by contradiction. As for all $j \leq p$, we are in one of the three following cases,
(a) i_{j-1}, i_{j+1} are smaller than i_j, and

$$|\Delta \mathsf{v}_j| = \left| \mathsf{v}_{q_{i_j}}(\tau_{i_j}^-) - \mathsf{v}_{q'_{i_j}}(\tau_{i_j}^-) \right|,$$

(b) i_{j-1}, i_{j+1} are bigger than i_j, and

$$|\Delta \mathsf{v}_j| = \left| \mathsf{v}_{q_{i_j}}(\tau_{i_j}^+) - \mathsf{v}_{q'_{i_j}}(\tau_{i_j}^+) \right|,$$

(c) $i_{j-1} < i_j < i_{j+1}$ or $i_{j-1} < i_j < i_{j+1}$.
As the difference of velocities is conserved by the scattering map and we are in $\mathscr{E}_1$, the cases (a) and (b) are impossible. Using that $i_0 < i_1$, we deduce that $\tau_{i_0} < \tau_{i_1} < \cdots < \tau_{i_p} < \tau_{\mathrm{rec}}$. In the same way, $\tau_{i'_0} < \tau_{i'_1} < \cdots < \tau_{i'_{p'}} < \tau_{\mathrm{rec}}$.

Using triangular inequality,

$$|w_f - w_0| = \left| \sum_{j=1}^{p} \Delta \mathsf{v}_j - \sum_{j=1}^{p'} \Delta \mathsf{v}'_j \right| \leq \frac{\varepsilon^{c_1}}{3t}.$$

Considering the $s \in [\tau_{i_0+1}, t]$ such that the minimum of (B.11) is reached, we have

$$\begin{aligned} |(\tau_{i_0+1} - \tau_{i_0}) w_0| &\leq |(\tau_{i_0} - s) w_0| \\ &\leq \varepsilon^{1-c_2} + \left(\sum_{j=1}^{p} |\Delta \mathsf{v}_j| + \sum_{j=1}^{p'} |\Delta \mathsf{v}'_j| \right) t + s|w_f - w_0| \\ &\leq \varepsilon^{1-c_2} + \frac{2\varepsilon^{c_1}}{3}, \end{aligned}$$

which is impossible since we are in $\mathscr{E}_1$. □

As $|\Delta \mathsf{v}_{\ell'} \times \hat{w}_{\mathrm{f}}| \leq \frac{\varepsilon^{c_3}}{n_k}$, (else it would have been treated in Point (1))

$$\left| \frac{\Delta \mathsf{v}_{\ell_0}}{|\Delta \mathsf{v}_{\ell_0}|} \times \hat{w}_{\mathrm{f}} \right| \leq \frac{t\varepsilon^{c_3}}{n_k |\Delta \mathsf{v}_{\ell_0}|} \leq t\varepsilon^{c_3 - c_1} \text{ and } |w_0 \times \hat{w}_f| \leq \varepsilon^{c_3}.$$

Finally, we obtain

$$\left| \frac{\Delta \mathsf{v}_{\ell_0}}{|\Delta \mathsf{v}_{\ell_0}|} \times w_0 \right| \leq \varepsilon^{c_3} + t\varepsilon^{c_3 - c_1}. \tag{B.12}$$

Definition B.1. *For $\omega \in \mathbb{R}^d$ and $\nu \in \mathbb{S}^{d-1}$, we denote (using ξ_α defined in Definition 2.1)*

$$(\tfrac{w'}{2}, -\tfrac{w'}{2}, \nu') = \xi_\alpha((w/2, -w/2, \nu)).$$

We define

$$\zeta_1(w,\nu) := w,\ \zeta_2(w,\nu) := w',\ \zeta_3(w,\nu) := \frac{w - w'}{2} \text{ and } \zeta_4(w,\nu) := \frac{w + w'}{2}. \tag{B.13}$$

Note that $\Delta \mathsf{v}_{\ell_0}$ is equal one of the $(\pm\zeta_i(\mathsf{v}_{q_{\ell_0}}(\tau_{\ell_0}^-) - \mathsf{v}_{q'_{\ell_0}}(\tau_{\ell_0}^-), \nu_{\ell_0})_{i \leq 4}$.
We conclude that

$$\begin{aligned}
\int_{\mathfrak{T}_{n_k}} \mathrm{d}\tau_{n_k} \mathbb{1}_{\text{contact}} \leq \frac{C^{n_k} t^{n_{k-1}} \theta^{(n_k - n_{k-1} - 1)_+}}{n_k^{n_k}} \Bigg(& \varepsilon^{1-c_2-c_3} + \\
& + \sum_i \mathbb{1}_{|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{q_i}(\tau_i^-)| \leq \varepsilon^{c_1}} \tfrac{\varepsilon^{c_1}}{|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{q_i}(\tau_i^-)|} + \mathbb{1}_{\frac{|(\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)) \times \nu_i|^2}{|\mathsf{v}_{q_i}(\tau_i^-) - \mathsf{v}_{\bar{q}_i}(\tau_i^-)|} \leq \varepsilon^{c_2}} \\
& + \sum_{\substack{y_0 \in \mathbb{Z}^d \\ |y_0| \leq t n_k \mathbb{V}}} \sum_{\substack{i \leq n_k \\ (q,q')}} \mathbb{1}_{|(\mathsf{v}_q(\tau_j^-) - \mathsf{v}_{q'}(\tau_j^-)) \times y_0| \leq 2t\mathbb{V}\varepsilon^{1-c_2}} \\
& + \sum_{i,j} \sum_{\ell=1}^{4} \mathbb{1}_{\left| \frac{\zeta_\ell(\mathsf{v}_{q_j}(\tau_j^-) - \mathsf{v}_{q'_j}(\tau_j^-), \nu_j)}{|\zeta_\ell(\mathsf{v}_{q_j}(\tau_j^-) - \mathsf{v}_{q'_j}(\tau_j^-), \nu_j)|} \times (\mathsf{v}_{q_i}(\tau_i^+) - \mathsf{v}_{q'_i}(\tau_i^+)) \right| \leq t\varepsilon^{c_3 - c_1}} \Bigg).
\end{aligned} \tag{B.14}$$

We have to integrate now with respect to $(\nu_{[n_k-1]}, V_{n_k})$. As in the proof of (6.10), we use the applications Ξ_T^i defined by (6.12).
We recall that

$$\Xi_T^{i+1} \Xi_T^i \cdots \Xi_T^1 (\nu_{[n_k-1]}, V_{n_k}) = (\tilde{\nu}_{[n_k-1]}, \tilde{V}_{n_k} = \mathsf{V}_{n_k}(\tau_i^+))$$

and that the Jacobian of this application is 1. We can integrate with respect to $\mathsf{v}_q(\tau_i^-) - \mathsf{v}_q(\tau_i^-)$.

We treat only the last singularity; the other ones can be estimated in the same way. If we remove the edges (q_j, q'_j) from T, either q_j or q'_j is not in the connected component of $\{q_i, q'_i\}$. Without loss of generality, we suppose that it is q_j. We denote ω the connected components of q_j in $T \setminus \{(q_j, q'_j, s_j)\}$. Before the collision j, the particles of ω are independent of the other ones, and as before, we can construct an application of Jacobian 1

$$\bar{\Xi} : (\nu_{[n_k-1]}, V_{n_k}) \mapsto (\bar{\nu}, \mathsf{V}_{\omega^c})(\tau_i^+), \mathsf{V}_\omega(\tau_j^-)).$$

In addition,

$$\begin{aligned}
|V_{n_k}|^2 &= \frac{|\mathsf{V}_{n_k}(\tau_i^+)|^2}{2} + \frac{|\mathsf{V}_{n_k}(\tau_j^-)|^2}{2} \\
&\geq \frac{|\mathsf{V}_{\omega^c}(\tau_i^+)|^2}{2} + \frac{|\mathsf{v}_{q_j}(\tau_j^-)|^2 + |\mathsf{v}_{q_j}(\tau_j^-)|^2}{2} + \frac{|\mathsf{V}_{\omega \setminus \{q_j\}}(\tau_j^-)|^2}{2} \\
&\geq \frac{|\mathsf{V}_{\omega^c}(\tau_i^+)|^2}{2} + \frac{|\mathsf{v}_{q_j}(\tau_j^-) - \mathsf{v}_{q_j}(\tau_j^-)|^2}{4} + \frac{|\mathsf{V}_{\omega \setminus \{q_j\}}(\tau_j^-)|^2}{2},
\end{aligned}$$

and denoting $w := \mathsf{v}_{q_j}(\tau_j^-) - \mathsf{v}_{q_j}(\tau_j^-)$, we can integrate with respect to the velocities

$$\sum_T \int_{\mathfrak{G}_T^0} \mathbb{1}_{\left| \frac{\zeta_\ell(\mathsf{v}_{q_j}(\tau_j^-)-\mathsf{v}_{q_j'}(\tau_j^-),\nu_j)}{|\zeta_\ell(\mathsf{v}_{q_j}(\tau_j^-)-\mathsf{v}_{q_j'}(\tau_j^-),\nu_j)|} \times (\mathsf{v}_{q_i}(\tau_i^+)-\mathsf{v}_{q_i'}(\tau_i^+)) \right| \le \varepsilon^{1/4}} \Lambda_T(V_{n_K}, \nu_{[n_K-1]}) M^{\otimes n_k} \,\mathrm{d}\nu_{[n_k-1]} \,\mathrm{d}V_{n_k}$$

$$\le \int \mathbb{1}_{\left| \frac{\zeta_\ell(w,\nu_j)}{|\zeta_\ell(w,\nu_j)|} \times (v_{q_i}-v_{q_i'}) \right| \le \varepsilon^{1/4}} \frac{(C(1+|V_{n_k}|^2))^{n_k} e^{-\frac{|V_{[n_k]\setminus\{q_j\}}|^2}{4} - \frac{|w|^2}{8}}}{(2\pi)^{\frac{n_k d}{2}}} \,\mathrm{d}\nu_{[n_k-1]} \,\mathrm{d}\tilde{V}_{[n_k]} \,\mathrm{d}w$$

$$\le n_k^{2n_k} C^{n_k} \varepsilon^{c_3-c_1}.$$

We have first integrate with respect to $(v_{q_i}, c_{q_i'})$ and then the (w, ν_j). Finally, we obtain

$$\frac{1}{(\mu\mathfrak{d})^{n_k-1}} \sum_T \int_{\mathfrak{T}_{n_k}\times\mathfrak{G}_T^0} \mathbb{1}_{\mathrm{reco}} \prod_{i=1}^{n_k-1} |(\mathsf{v}_{q_i}(\tau_i) - \mathsf{v}_{q_i'}(\tau_i)) \cdot \nu_i| \,\mathrm{d}\nu_i \,\mathrm{d}\tau_i M^{\otimes n_k} dV_{n_k}$$

$$\le \frac{(n_k C)^{n_k} t^{n_{k-1}} \theta^{(n_k-n_{k-1}-1)_+}}{(\mu\mathfrak{d})^{n_k-1}} \left(\varepsilon^{\frac{c_2}{2}} + \varepsilon^{c_3-c_1} + \varepsilon^{c_1} + \varepsilon^{1-c_2} + \varepsilon^{1-c_2-c_3} \right). \tag{B.15}$$

We choose $c_1 = \frac{1}{5}$ and $c_2 = c_3 = \frac{2}{5}$. This concludes the proof. □

B.3. **Proof of** (3.6)**.**

Lemma B.3. *For $r \le 2\gamma$,*

$$\int \mathbb{1}_{\substack{Z_r \text{ form a} \\ \text{possible cluster}}} M^{\otimes r} \,\mathrm{d}Z_r \le C_\gamma \mu^{-r+1} \delta^{r-1}. \tag{B.16}$$

Proof. First choose a family $\omega_1, \cdots, \omega_p$ of subset covering $[r]$, and $(\underline{\lambda}_i)_{i\le p} = (\lambda_i^1, \cdots, \lambda_i^{l_i})_{i\le p}$ a family of partitions of ω_i. As n is bounded, there are a finite number of possible $((\omega_i)_i, (\underline{\lambda}_i)_i)$. We construct the graph $\mathscr{G}$ as the merge of the collision graph of $\mathsf{Z}(\tau, Z_{\omega_i}, \underline{\lambda}_i)$ on $[0,\delta]$, and we extract T the clustering tree (there are less than $(2\gamma)^{2\gamma}$ possible clustering trees). We can then adapt the proof of (5.3) (where we treated only two pseudotrajectories), and we obtain the expected result. □

B.4. **Proof of** (8.10)**.** We recall that $\mathscr{O}_r \subset \mathbb{D}^r$ is the set

$$\mathscr{O}_r := \Big\{ Z_r \in \mathbb{D}^r,\ \exists (\lambda_1, \cdots, \lambda_{\mathfrak{l}}),\ \text{the collision graph of } \mathsf{Z}_r(\cdot, Z_z, (\lambda_1, \cdots, \lambda_{\mathfrak{l}})) \text{ on } [0,\delta] \text{ is}$$
$$\text{connected and the pseudotrajectory has a collision or a multiple interaction} \Big\}.$$

Proposition B.4. *There exists a positive constant C_r depending only on the dimension and the number of particles r such that*

$$\int_{\mathbb{T}^{r-1}\times B_r(\mathbb{V})} \mathbb{1}_{\mathscr{O}_r}(Z_r) \frac{e^{-\mathscr{H}_r}}{(2\pi)^{\frac{rd}{2}}} \,\mathrm{d}X_{2,r} \,\mathrm{d}V_r \le \frac{C_r}{(\mu\mathfrak{d})^{r-1}} \delta^2 \theta^{r-2} \varepsilon^{\mathfrak{1}/\mathfrak{12}}, \tag{B.17}$$

where $B_r(\mathbb{V})$ is the ball of radius $\mathbb{V}$ in dimension rd (we use that $\delta = \varepsilon^{1/12}$).

Proof. For $r = 2$, we have

$$\int_{\mathbb{D}^2} \mathbb{1}_{\mathscr{O}_2}(Z_2) \frac{e^{-\mathscr{H}_2}}{(2\pi)^d} \,\mathrm{d}x_2 \,\mathrm{d}V_2 = C\varepsilon^d \le \frac{C\delta^2\varepsilon^{1/12}}{\mu\mathfrak{d}}$$

as $\delta^2\varepsilon^{1/12} = \varepsilon^{3/12} \le \varepsilon$.

Fix parameters $(\lambda_1, \cdots \lambda_{\mathfrak{l}})$ and denote $\mathscr{R}_{(\lambda_1,\cdots\lambda_{\mathfrak{l}})} \subset \mathbb{D}^r$ the set of initial configuration such that the pseudotrajectory has a connected collision graph and a pathological collision.

As we suppose that the collision graph is connected, we can construct a clustering tree $T := (q_i, \bar{q}_i)_{i\le r-1}$ as in the proof of Proposition 5.2. We define τ_{path} the time of the first pathological collision. The corresponding collision can either create a loop in the collision graph or be a clustering multiple encounter
.

The first case can be treated as a recollision, which is already done in the proof of Proposition 5.2 and in the preceding section, we have

$$\int_{\mathbb{T}^{r-1}\times B_r(\mathbb{V})} \mathbb{1}_{\mathscr{R}^{\mathrm{reco}}_{(\lambda_1,\cdots\lambda_{\mathfrak{l}})}} \frac{e^{-\mathscr{H}_r}}{(2\pi)^{\frac{rd}{2}}} \,\mathrm{d}X_{2,r} \,\mathrm{d}V_r \le \frac{C_r\delta^{r-2}}{(\mu\mathfrak{d})^{r-1}} \varepsilon^{1/5} \le \frac{C_r}{(\mu\mathfrak{d})^{r-1}} \delta^2 \theta^{r-2} \varepsilon^{\mathfrak{1}/\mathfrak{12}}.$$

In the second case, there are two clustering collisions $j < \tilde{j}$ such that $\{q_j, \bar{q}_j\} \cap \{q_{\tilde{j}}, \bar{q}_{\tilde{j}}\}$ and

$$\forall \tau \in (\tau_j, \tau_{\tilde{j}}),\ |\mathsf{x}_{q_j}(\tau) - \mathsf{x}_{\bar{q}_j}(\tau)| < \varepsilon.$$

Two particles (q_j, q_j') can stay at a distance shorter than ε on a time interval shorter than

$$1 \wedge \frac{\varepsilon |v_{q_j}(\tau_j) - v_{\bar{q}_j}(\tau_j)|}{|(v_{q_j}(\tau_j) - v_{\bar{q}_j}(\tau_j)) \wedge (x_{q_j}(\tau_j) - x_{\bar{q}_j}(\tau_j))|^2}$$

(whose integral is bounded by $\varepsilon^{\frac{1}{2}}$). Hence, using the same strategy than in the proof of Proposition 5.2:

$$\int_{\mathbb{T}^{r-1}\times B_r(\mathbb{V})} \mathbb{1}_{\mathscr{R}^{\mathrm{mult}}_{(\lambda_1,\cdots\lambda_1)}} \frac{e^{-\mathscr{H}_r}}{(2\pi)^{\frac{rd}{2}}}\,\mathrm{d}X_{2,r}\,\mathrm{d}V_r \le \frac{C_r\delta^{r-1}}{(\mu\mathfrak{d})^{r-2}}\varepsilon^{1/2} \le \frac{C_r}{(\mu\mathfrak{d})^{r-1}}\delta^2\theta^{r-2}\varepsilon^{1/12}.$$

Summing on all the possible $(\lambda_1, \cdots \lambda_1)$, we obtain the expected result. □

Proof of (8.10). We have now to prove Estimation (8.10):

$$\int_{\mathbb{T}^{r-1}\times(\mathbb{R}^d)^r} \mathbb{1}_{\mathscr{O}_\varpi} \mathbb{1}_{\substack{Z_r \text{ form a}\\ \text{possible cluster}}} \frac{e^{-\frac{1}{2}\mathscr{H}_r(Z_r)}}{(2\pi)^{dr/2}}\,\mathrm{d}X_{2,r}\,\mathrm{d}V_r \le \frac{C_r}{(\mu\mathfrak{d})^{r-1}}\delta^2\theta^{r-2}\varepsilon^{1/12},$$

Without loss of generality, we suppose that $1 \in \varpi$.

Fix the family $\omega_1, \cdots, \omega_p$ of subset covering $[n]$, and $(\underline{\lambda}_i)_{i\le p} = (\lambda_i^1, \cdots, \lambda_i^{l_i})_{i\le p}$ a family of partition of ω_i such that the union of the collision graph associated to parameters $((\lambda_i^j)_j)_i$ is connected.

We begin by fix Z_ϖ and sum the clustering of the particle in $[n] \setminus \varpi$

$$\int_{\mathbb{T}^{n-1}\times(\mathbb{R}^d)^n} \mathbb{1}_{\mathscr{O}_\varpi} \mathbb{1}_{\substack{Z_r \text{ form a}\\ \text{possible cluster}}} \frac{e^{-\frac{1}{2}\mathscr{H}_n(Z_n)}}{(2\pi)^{dn/2}}\,\mathrm{d}X_\varpi\,\mathrm{d}V_\varpi \le \mathbb{1}_{\mathscr{O}_\varpi} \frac{e^{-\frac{1}{2}\mathscr{H}_{|\varpi|}(Z_\varpi)}}{(2\pi)^{d|\varpi|/2}} \frac{C_n\delta^{n-|\varpi|}}{(\mu\mathfrak{d})^{|\varpi|}}.$$

Then, we integrate with respect to $\mathrm{d}X_{\varpi\setminus\{1\}}\,\mathrm{d}V_\varpi$. □

Acknowledgment: The author thanks Sergio Simonella and Raphael Winter for their numerous suggestions as well as for the misprint corrections which helped to improve the paper. He also thanks Laure Saint-Raymond for stimulating and fruitful discussions. The author acknowledges financial support from the European Union (ERC, PASTIS, Grant Agreement n°101075879).[10]

References

[Ayi17] Nathalie Ayi. From newton's law to the linear boltzmann equation without cut-off. *Communications in Mathematical Physics*, 350(3):1219–1274, Mar 2017.

[BGSR16] Thierry Bodineau, Isabelle Gallagher, and Laure Saint-Raymond. The Brownian motion as the limit of a deterministic system of hard-spheres. *Invent. Math.*, 203(2):493–553, 2016.

[BGSR17] Thierry Bodineau, Isabelle Gallagher, and Laure Saint-Raymond. From hard sphere dynamics to the Stokes-Fourier equations: an L^2 analysis of the Boltzmann-Grad limit. *Ann. PDE*, 3(1):118, 2017. Id/No 2.

[BGSRS18] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. One-sided convergence in the Boltzmann-Grad limit. *Ann. Fac. Sci. Toulouse, Math. (6)*, 27(5):985–1022, 2018.

[BGSRS20] T. Bodineau, I. Gallagher, L. Saint-Raymond, and S. Simonella. Statistical dynamics of hard sphere gas: fluctutating Boltzmann equation and large deviation. 2020.

[BGSRS22] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. Cluster expansion for a dilute hard sphere gas dynamics. *J. Math. Phys.*, 63(7):26, 2022. Id/No 073301.

[BGSRS23a] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Simonella Sergio. Statistical dynamics of a hard sphere gas: fluctuating boltzmann equation and large deviations. *Annals of Mathematics*, 198(3):1047–1201, 2023.

[BGSRS23b] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. Long-time correlations for a hard-sphere gas at equilibrium. *Commun. Pure Appl. Math.*, 76(12):3852–3911, 2023.

[BGSRS24] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. Long-time derivation at equilibrium of the fluctuating Boltzmann equation. *Ann. Probab.*, 52(1):217–295, 2024.

[Bol96] Ludwig Boltzmann. Vorlesungen über Gastheorie. 1. Teil. Theorie der Gase mit ein-atomigen Molecülen, deren Dimensionen gegen die mittlere Weglänge verschwinden. Leipzig: J. A. Barth. viii, 204 S. (1896)., 1896.

[BPS13] A. Boblylev, M. Pulvirenti, and C. Saffirio. From Particle Systems to the Landau Equation: A Consistency Result. *Commun. Math. Phys.*, 319(3):683–702, 2013.

[Cat18] Nicolo' Catapano. The rigorous derivation of the linear Landau equation from a particle system in a weak-coupling limit. *Kinet. Relat. Models*, 11(3):647–695, 2018.

[CIP94] Carlo Cercignani, Reinhard Illner, and Mario Pulvirenti. *The mathematical theory of dilute gases*, volume 106 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1994.

[10] Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

[Den18] Ryan Denlinger. The propagation of chaos for a rarefied gas of hard spheres in the whole space. *Arch. Ration. Mech. Anal.*, 229(2):885–952, 2018.

[DHM24] Yu Deng, Zaher Hani, and Xiao Ma. Long time derivation of the Boltzmann equation from hard sphere dynamics. Preprint, arXiv:2408.07818 [math.AP] (2024), 2024.

[DP99] L. Desvillettes and M. Pulvirenti. The linear boltzmann equation for long-range forces: a derivation from particle systems. *Math. Models Methods Appl. Sci.*, 09(08):1123–1145, 1999. Publisher: World Scientific Publishing Co.

[DR01] L. Desvillettes and V. Ricci. A rigorous derivation of a linear kinetic equation of Fokker–Planck type in the limit of grazing collisions. *Journal of Statistical Physics*, 104(5):1173–1189, 2001.

[Due21] Mitia Duerinckx. On the size of chaos via Glauber calculus in the classical mean-field dynamics. *Commun. Math. Phys.*, 382(1):613–653, 2021.

[Fou24] Florent Fougères. On the derivation of the linear Boltzmann equation from the nonideal Rayleigh gas. Adaptive pruning and improvement of the convergence rate. *J. Stat. Phys.*, 191(10):16, 2024. Id/No 136.

[Gal69] Giovanni Gallavotti. Divergences and the approach to equilibrium in the lorentz and the wind-tree models. *Physical Review*, 185(1):308, 1969.

[Gou97] T. Goudon. On Boltzmann equations and Fokker-Planck asymptotics: Influence of grazing collisions. *J Stat Phys*, 89(3):751, 1997.

[Gra49] Harold Grad. On the kinetic theory of rarefied gases. *Comm. Pure Appl. Math.*, 2:331–407, 1949.

[GSRT13] Isabelle Gallagher, Laure Saint-Raymond, and Benjamin Texier. *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013.

[IP89] R. Illner and M. Pulvirenti. Global validity of the Boltzmann equation for two- and three-dimensional rare gas in vacuum. Erratum and improved result: "Global validity of the Boltzmann equation for a two-dimensional rare gas in vacuum" [Comm. Math. Phys. **105** (1986), no. 2, 189–203; MR0849204 (88d:82061)] and "Global validity of the Boltzmann equation for a three-dimensional rare gas in vacuum" [ibid. **113** (1987), no. 1, 79–85; MR0918406 (89b:82052)] by Pulvirenti. *Comm. Math. Phys.*, 121(1):143–146, 1989.

[Jan] S. Jansen. Gibbsian point process. https://www.mathematik.uni-muenchen.de/jansen/gibbspp.pdf.

[Kin75] Francis Gordon King. BBGKY hierarchy for positive potentials. Technical report, University of California, Berkeley, United States, 1975. INIS Reference Number: 8284567.

[Lan36] L. Landau. Die kinetische Gleichung für den Fall Coulombscher Wechselwirkung. *Phys. Zs. Sow. Union*, 10(154):163–170, 1936.

[Lan75] Oscar E. Lanford, III. Time evolution of large classical systems. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, pages 1–111. Lecture Notes in Phys., Vol. 38. 1975.

[LB23] Corentin Le Bihan. *Effets de bord et comportement en temps long en théorie cinétique collisionnelle*. PhD thesis, Ecole normale supérieure de lyon-ENS LYON, 2023.

[LB25] Corentin Le Bihan. Long time validity of the linearized Boltzmann equation for hard spheres: a proof without billiard theory. *Arch. Ration. Mech. Anal.*, 249(4):72, 2025. Id/No 40.

[LBW22] Corentin Le Bihan and Raphael Winter. The grazing collisions limit from the linearized Boltzmann equation to the landau equation for short-range potentials. *Kinetic and Related Models*, 2022. https://arxiv.org/pdf/2208.14918.pdf.

[Max03] JC Maxwell. On the dynamical theory of gases, inthe kinetic theory of gases: An anthology of classic papers with historical commentary, sg brush, editor, 2003.

[NSV18] Alessia Nota, Sergio Simonella, and Juan J. L. Velázquez. On the theory of Lorentz gases with long range interactions. *Rev. Math. Phys.*, 30(3):62, 2018. Id/No 1850007.

[NVW21] Alessia Nota, Juan J. L. Velázquez, and Raphael Winter. Interacting particle systems with long-range interactions: scaling limits and kinetic equations. *Atti Accad. Naz. Lincei, Cl. Sci. Fis. Mat. Nat., IX. Ser., Rend. Lincei, Mat. Appl.*, 32(2):335–377, 2021.

[NVW22] Alessia Nota, Juan J. L. Velázquez, and Raphael Winter. Interacting particle systems with long-range interactions: approximation by tagged particles in random fields. *Atti Accad. Naz. Lincei, Cl. Sci. Fis. Mat. Nat., IX. Ser., Rend. Lincei, Mat. Appl.*, 33(2):439–506, 2022.

[Pen63] Oliver Penrose. Convergence of fugacity expansions for fluids and lattice gases. *Journal of Mathematical Physics*, 4(10):1312–1320, 1963.

[PS15] M. Pulvirenti and S. Simonella. On the evolution of the empirical measure for the hard-sphere dynamics. *Bull. Inst. Math. Acad. Sin. (N.S.)*, 10(2):171–204, 2015.

[PS21] Mario Pulvirenti and Sergio Simonella. A brief introduction to the scaling limits and effective equations in kinetic theory. *Trails in Kinetic Theory: Foundational Aspects and Numerical Methods*, pages 183–208, 2021.

[PSS14] M. Pulvirenti, C. Saffirio, and S. Simonella. On the validity of the Boltzmann equation for short range potentials. *Rev. Math. Phys.*, 26(2):1450001, 64, 2014.

[PSW16] Robert I. A. Patterson, Sergio Simonella, and Wolfgang Wagner. Kinetic theory of cluster dynamics. *Physica D*, 335:26–32, 2016.

[Sin72] Ya. G. Sinai. Construction of dynamics in one-dimensional systems of statistical mechanics. *Theoretical and Mathematical Physics*, 11:487 – 494, 1972.

[Spo81] Herbert Spohn. Fluctuations around the Boltzmann equation. *J. Statist. Phys.*, 26(2):285–305, 1981.

[Spo83] Herbert Spohn. Fluctuation theory for the boltzmann equation. *Nonequilibrium Phenomena I: The Boltzmann Equation*, pages 225–251, 1983.

[SW25] Sergio Simonella and Raphael Winter. Pointwise decay of cumulants in chaotic states at low density. *Markov processes and related fields*, 2025.

[vBLLS80] Henk van Beijeren, O. E. III Lanford, J. L. Lebowitz, and H. Spohn. Equilibrium time correlation functions in the low-density limit. *J. Stat. Phys.*, 22:237–257, 1980.

[Vil98] Cédric Villani. On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Ration. Mech. Anal.*, 143(3):273–307, 1998.

[Win21] R. Winter. Convergence to the Landau equation from the truncated BBGKY hierarchy in the weak-coupling limit. *Journal of Differential Equations*, 283:1–36, 2021.