VALIDITY AND FAILURE OF THE BOLTZMANN APPROXIMATION OF KINETIC ANNIHILATION

KARSTEN MATTHIES AND FLORIAN THEIL

Abstract. This paper introduces a new method to show the validity of a continuum description for the deterministic dynamics of many interacting particles. Here the many particle evolution is analyzed for a hard sphere flow with the addition that after a collision the collided particles are removed from the system. We consider random initial configurations which are drawn from a Poisson point process with spatially homogeneous velocity density $f_0(v)$. Assuming that the moments of order less than three of f_0 are finite and no mass is concentrated on lines, the homogeneous Boltzmann equation without gain term is derived for arbitrary long times in the Boltzmann-Grad scaling. A key element is a characterization of the many particle flow by a hierarchy of trees which encode the possible collisions. The occurring trees are shown to have favorable properties with a high probability, allowing to restrict the analysis to a finite number of interacting particles, enabling us to extract a single-body distribution. A counter-example is given for a concentrated initial density f_0 even to short-term validity.

Keywords: Boltzmann equation, Boltzmann-Grad limit, validity, kinetic annihilation, deterministic dynamics, random initial data

Classification: 82C40, 76P05, 82C22, 60K35,

The derivation of the continuum models of mathematical physics from atomistic descriptions is a longstanding and fundamental problem. This includes e.g. the emergence of irreversible macroscopic behavior generated by deterministic reversible Hamiltonian micro-evolution.

An illustration of this question is provided by deterministic hard ball dynamics with random initial states. For high particle numbers and suitably scaled diameters it is expected that the time-evolution of the density is close to the solution of the Boltzmann equation

$$\partial_t f + v \cdot \partial_u f = \int_{\mathbb{R}^d \times S^{d-1}} (f(u,\tilde{v})f(u,\tilde{v}') - f(u,v)f(u,v'))((v-v')\cdot\nu)_+ \,\mathrm{d}v' \,\mathrm{d}\nu, \tag{1}$$

where $g_+ = \max(g,0)$ is the positive part, $\tilde{v}, \tilde{v}'$ are obtained from v, v' by exchanging the respective components of v and v' in direction ν, that is

$$\tilde{v} = v + (v'-v)\cdot\nu\,\nu, \quad \tilde{v}' = v' + (v-v')\cdot\nu\,\nu,$$

and $f_t(u,v)$ is the density of presence at time t of particles at locations u with velocity v, see [Spo91].

An important concept which sheds some light on the connection between the Boltzmann equation and hard ball dynamics is the propagation of chaos. Though the distribution $p_N(u_1, v_1 \ldots, u_N, v_N, t)$ of N particles loses its product structure for nonzero time t, the marginal distribution of the first k particles should be very close to a product measure when the total number of particles N is large. A classical method to establish propagation of chaos is to express the evolution of k-particle marginals in terms of the $k+1$-particle marginals. This strategy is implemented in the BBGKY hierarchy. The weakness of this approach consists in the fact that establishing convergence of the resulting series is hard in many cases. O. Lanford succeeded in proving that in the case of hard ball dynamics the series that corresponds to the BBGKY hierarchy converges for small times to a solution of the Boltzmann equation [Lan75]. Unfortunately it cannot be shown that the time interval where the series is known to converge is larger than a small fraction of the mean free flight time, regardless of the initial data. This problem was partially overcome by [IP89] who managed to obtain a global result if the positions are in $\mathbb{R}^d$ and the initial density is sufficiently small. Other related results can be found in [Gal70, Lan75, Spo78, BBS83, Spo91, CIP94] and references therein. However, currently there is no result which covers the case where both data and time are large. It is arguable that the justification of the Boltzmann equation (1) as a scaling limit of deterministic evolution constitutes a part of Hilbert's sixth problem [Hil00].

In [LN80] the same strategy is applied to the simpler problem of coagulation. Here the spheres move along Brownian paths and two intact spheres annihilate each other if the distance between the centers drops below a. Although the series generated by the BBGKY hierarchy does not converge globally in time, Lang and Nguyen were able to give a rigorous justification of the corresponding Boltzmann equation by restarting the procedure at small positive time.

In this paper we consider kinetic annihilation, another simplification of hard ball dynamics which keeps two central features of the original evolution: The initial state is random, the evolution is deterministic. We assume that the initial configuration ω is a finite subset of the phase space $\mathbb{T}^d \times \mathbb{R}^d$ ($\mathbb{T}^d$ is the unit torus) and is drawn from a Poisson point process with some intensity $\mu \in M_+(\mathbb{T}^d \times \mathbb{R}^d)$. As long as they are intact the centers of the spheres move along straight lines with constant velocity. When the centers of two spheres, which are still intact, come within distance a, then both spheres are destroyed. Another term for this type of evolution is "ballistic annihilation".

We will consider the asymptotic behavior of the system in the limit where the diameter a of the particles tends to 0 and the total intensity $n = \mu(\mathbb{T}^d \times \mathbb{R}^d)$ is linked to a by the Boltzmann-Grad relation

$$na^{d-1} = 1. \tag{2}$$

The central question in this paper is whether for small values of a the many-body evolution can be described by the gainless Boltzmann equation

$$\partial_t f + v \cdot \partial_u f = Q_-[f,f], \tag{3}$$

where $f(u,v)$ is the distribution function for $(u,v) \in \mathbb{T}^d \times \mathbb{R}^d$; the expression $Q_-[f,g](v) = -\kappa_d f(v) \int_{\mathbb{R}^d} \mathrm{d}g(v')\, |v - v'|$ is the loss term of the hard-sphere collision kernel of the Boltzmann equation (1) and κ_d is the volume of the $d-1$ dimensional unit-ball. For the sake of simplicity we will restrict ourselves to the case where the initial density f_0 does not depend on u, in this case the transport term $v \cdot \partial_u f$ in eq. (3) vanishes and $f_t(u,v) = f_t(v)$. We will establish the validity of the Boltzmann equation (3) in the following, probabilistic sense: Let $(u(t), v(t))$ be the position and velocity of a tagged particle at time t, then for $A \subset \mathbb{T}^d \times \mathbb{R}^d$ Borel.

$$\begin{aligned} &\lim_{a \to 0} \mathrm{Prob}((u(t), v(t)) \in A \text{ and the particle is intact at time } t) \\ =&\frac{1}{f_0(\mathbb{R}^d)} \int_A \mathrm{d}u\, \mathrm{d}f_t(v). \end{aligned} \tag{4}$$

Since the distribution of the N particles is invariant under permutation it is irrelevant which particle index we use to define the validity. Following standard proofs of strong laws of large numbers, see e.g. [Dur], simple bounds on correlations which are beyond the scope of this paper, can be used to deduce that the validity of the Boltzmann equation in the sense of eq. (4) implies that the solution f can also be interpreted as a density, i.e.

$$\lim_{a\to 0} \mathrm{Prob}\Bigg(\Bigg|\tfrac{1}{n}\#\{i \mid (u_i(t), v_i(t)) \in A \text{ and particle } i \text{ is intact at time } t\} \\ - \int_A \mathrm{d}u\, \mathrm{d}f_t(v)\}\Bigg| > \varepsilon\Bigg) = 0$$

for all $\varepsilon > 0$ and all $A \subset \mathbb{T}^d \times \mathbb{R}^d$ Borel.

Kinetic annihilation dynamics can be used to model growth and coarsening of surfaces, see [KS88], and has been studied extensively in the physics literature, see [EF85, Pia95, DFPR95, PTD02, CDPTW03].

The main result of this paper is a rigorous proof that the gainless Boltzmann equation (3) is valid in the sense of eq. (4), provided that $f_0 \in M_+(\mathbb{T}^d \times \mathbb{R}^d)$ is homogeneous (i.e. $f_0(u, v) = f_0(v)$ for all u), has finite total mass and kinetic energy

$$\int_{\mathbb{R}^d} (1 + |v|)^2 \,\mathrm{d}f_0(v) = K_{\mathrm{ini}} < \infty \tag{5}$$

and does not concentrate mass on single velocity directions, i.e.

$$\int_{\rho(v,\nu)} \mathrm{d}f_0(v') = 0 \text{ for all } v \in \mathbb{R}^d, \nu \in S^{d-1}, \tag{6}$$

where $\rho(v, \nu) = v + \mathbb{R}\,\nu$ is a line.

The results were announced –without proof– in [MT08]. The assumption that f_0 is homogeneous will be dropped in a forthcoming publication. Bounds on the moments of f_0 are standard in the literature, but assumption (6) appears to be new. In Section 3 we will discuss an example which shows that this assumption cannot be dropped without losing the approximation property of the Boltzmann equation. We demonstrate that for arbitrarily short but finite times the limit of the empirical density is not consistent with the mean-field theory. This shows that further assumptions are needed in the informal justification of the gainless Boltzmann equation in [PTD02].

In the proof we insert an additional layer between the single-body densities and the N-body evolution: The probability distribution of trees which encode the collision history of the individual particles. A very similar approach has been used previously in [Sz91] in connection with coagulation dynamics. We introduce two separate distributions, the empirical tree distribution $\hat{P}$ which is extracted from the many body evolution and an idealized distribution P which is postulated and ignores correlations caused by rare events such as recollisions.

The main steps of the proof are concerned with clarifying the relation between trees, the single-body evolution and the many-body evolution:

(1) We construct explicit expressions for the empirical tree distribution $\hat{P}$ and the idealized distribution P.
(2) The convergence of the empirical distribution $\hat{P}$ to the limiting distribution P can be established within the set of good trees $\mathcal{G}$. Together with the proof that the complement of $\mathcal{G}$ is small, this amounts to establishing convergence of $\hat{P}$ to P in the total-variation sense.
(3) We show that f_t, the single body marginal of P, satisfies the gainless, homogeneous Boltzmann equation.

In section 4, we collect some proofs, which are not immediately needed in the understanding and the development of the concepts of this article. In section 5, we discuss conclusions, variants and extensions. An appendix with a list of frequently used notation is included.

1. Main result

On the atomistic level we consider N particles with initial values $(u_0(i), v_0(i)) \in \mathbb{T}^d \times \mathbb{R}^d$, $i = 1, \dots, N$, which evolve by force-free Newtonian dynamics

$$
\begin{aligned}
&u(i, t=0) = u_0(i),\ v(i, t=0) = v_0(i),\\
&\dot u(i,t) = v(i,t),\ \dot v(i,t) = 0.
\end{aligned}
\tag{7}
$$

For each $t \in [0,\infty)$, $i \in \{1,\dots,N\}$ there exists a unique scattering state $\beta^{(a)}(i,t) \in \{0,1\}$ which indicates whether the i-th particle has already collided ($\beta(i) = 0$) or not ($\beta(i) = 1$). We assume that particles that overlap initially do not collide, and obtain that β satisfies the implicit relation

$$
\beta^{(a)}(i,t) = \begin{cases} 1 & \text{if } \operatorname{dist}(z_i, z_{i'}, s) \geq a\beta^{(a)}(i', s) \text{ for all } s \in [0,t),\ i' \neq i,\\ 0 & \text{else} \end{cases}
\tag{8}
$$

with a modified distance function to ignore initial intersections

$$
\operatorname{dist}((u,v),(u',v'),t) = \begin{cases} 2a & \text{if } |u-u'|_{\mathbb{T}^d} < a \text{ and}\\ & \quad |u-u'+s(v-v')|_{\mathbb{T}^d} \leq a \text{ for all } s \in [0,t),\\ |u-u'+t(v-v')|_{\mathbb{T}^d} & \text{else.} \end{cases}
\tag{9}
$$

Here $|.|_{\mathbb{T}^d}$ is the distance on the torus, i.e. $|\tilde u|_{\mathbb{T}^d} = \inf_{k\in\mathbb{Z}^d} |\tilde u - k|_{\mathbb{R}^d}$. We are interested in the evolution of a tagged particle when the initial configuration is drawn according to a modified Poisson-point process. The modification accounts for the fact that the total number of particles in the system exceeds or equals 1. This concept is related to Palm measures of Poisson processes, see e.g. [Kal05, Sec 2.7].

Definition 1 (Tagged Poisson point processes)**.** *Let Ω be a locally compact metric space. The tagged particle z_1 is an independent random variable with law $\mu/\mu(\Omega)$. The random variable $\tilde z \in \bigcup_{r=0}^\infty \Omega^r$ forms a Poisson point process with density $\mu \in M_+(\Omega)$ (non-negative Radon measures, i.e. positive elements of $(C_c^0(\Omega))^*$) if*

$$
\operatorname{Prob}(\tilde z \in \Omega^r) = e^{-\mu(\Omega)} \frac{\mu(\Omega)^r}{r!}, \quad \operatorname{law}(\tilde z_i) = \mu/\mu(\Omega),
$$

and $\tilde z_1, \dots, \tilde z_r$ are independent. Now letting $N = r+1 \in \{1,2,\dots\}$, realizations of the tagged Poisson point process (tppp) are obtained by letting $z = (z_1,\dots,z_N) = (z_1, \tilde z)$, i.e. one obtains for symmetric $A \subset \bigcup_{N=1}^\infty \Omega^N$ that

$$
\operatorname{Prob}_{\text{tppp}}((z_1,\dots,z_N) \in A) = \frac{1}{\mu(\Omega) e^{\mu(\Omega)}} \sum_{N=1}^\infty \frac{1}{(N-1)!} \int_{A\cap\Omega^N} \mathrm{d}\mu(z_1)\dots\mathrm{d}\mu(z_N).
$$

Theorem 2. *(Validity of the gainless Boltzmann equation) Let the probability measure $f_0 \in PM(\mathbb{R}^d), d \geq 2$, be a momentum density that satisfies (5, 6). Let $\omega \subset \mathbb{T}^d \times \mathbb{R}^d$ be a realization of the tagged Poisson point process with intensity $\mu = n\,(\mathbf{1}_{\mathbb{T}^d} \otimes f_0)$, where $\mathbf{1}_{\mathbb{T}^d}$ is the standard Lebesgue measure restricted to the unit-torus, and n is determined by a and the Boltzmann-Grad scaling (2). If $N = \#\omega$ particles with initial values in ω evolve by (7), then for each $t \in [0,\infty)$*

$$
\lim_{a\to 0} \sup_{A\subset\mathbb{T}^d\times\mathbb{R}^d \text{ Borel}} \left| \operatorname{Prob}_{\text{tppp}}\Big(z(1,t) \in A \text{ and } \beta^{(a)}(1,t) = 1\Big) - \int_A \mathrm{d}u\,\mathrm{d}f_t(v) \right| = 0
\tag{10}
$$

where $f : [0,\infty) \to M_+(\mathbb{R}^d)$ is the unique solution of the homogeneous, gainless Boltzmann equation

$$\dot{f} = Q_-[f,f], \quad f_{t=0} = f_0, \tag{11}$$

with $Q_-[f,f](v) = -\int_{\mathbb{R}^d} \mathrm{d}f(v')\,\kappa_d|v-v'|\,f(v)$, and κ_d the volume of $d-1$ dimensional unit-ball, in particular $\kappa_2 = 2, \kappa_3 = \pi$.

Corollary 3. *The measures $\mathbf{1}_{\mathbb{T}^d} \otimes f_t$ and*

$$\mathrm{d}\hat{f}_t^{(a)}(u,v) = \mathrm{Prob}_{\mathrm{tppp}}(z(1,t) \in [u, u+\mathrm{d}v) \times [v, v+\mathrm{d}v) \,\mathrm{and}\, \beta^{(a)}(1,t) = 1)$$

are both absolutely continuous with respect to $\mathbf{1}_{\mathbb{T}^d} \otimes f_0$. Furthermore

$$\lim_{a\to 0} \hat{f}_t^{(a)} = \mathbf{1}_{\mathbb{T}^d} \otimes f_t \tag{12}$$

in the $L^1(\mathbf{1}_{\mathbb{T}^d} \otimes f_0)$ norm.

The proof of the theorem and the corollary can be found at the end of Section 2.

Remark 4. (1) *Note that the tagged Poisson point process is a symmetric point process. The motivation for working with this process is that the realizations of the tagged ppp without the tagged particle form a ppp and we obtain a very simple explicit formula for the distribution of trees, see (61), hence the complexity of the proof can be reduced. On the other hand, it seems that the formulae for the joint distribution of two trees are much more complicated, therefore we will only make statements which concern the law of a single, tagged particle.*

(2) *The assumption $\int_{\mathbb{R}^d} \mathrm{d}f_0(v) = 1$ is a standard normalization, but it is not necessary.*

(3) *Assumption (6) does not exclude the possibility that f_0 is concentrated on lower dimensional subsets, for example the uniform distribution on the sphere S^{d-1} is admissible, i.e. f_0 satisfies*

$$\int \varphi(v)\,\mathrm{d}f_0(v) := \frac{1}{\mathcal{H}^{d-1}(S^{d-1})} \int_{S^{d-1}} \varphi(v)\,\mathrm{d}\mathcal{H}^{d-1}(v), \tag{13}$$

for all testfunctions $\varphi \in C_c(\mathbb{T}^d \times \mathbb{R}^d)$, where $\mathcal{H}^d$ is the d-dimensional Hausdorff-measure.

(4) *We will analyze effects due to concentration by a Taylor expansion in time of f_t in Section 3.*

2. Proof of theorem 2

2.1. **The hierarchy of evolutions.** We replace the initial value problem (11) by an infinite system using general initial distribution without concentrations

$$\dot{f}_k = Q_-[f_{k-1}, f_k], \quad f_{t=0,k} = f_0. \tag{14}$$

Since Q_- is quadratic, for fixed k the integro-differential equation (14) is in fact linear and non-autonomous. The differential equation completely decouples in v and the equation for each v is a scalar linear non-autonomous ODE, which can be directly integrated to

$$f_{t,k} = \exp(-\textstyle\int_0^t L[f_{s,k-1}]\,\mathrm{d}s) f_0, \tag{15}$$

where $L[f](v) = \kappa_d \int \mathrm{d}f(v')\,|v-v'|$. We observe that $\mathrm{d}f_{t,k}(v)$ is absolutely continuous with respect to $\mathrm{d}f_0(v)$ due to the decoupling in v.

Lemma 5. *Let $f_0 \in M_{(1+|v|)^2}$ then f_k converges in $C^0_\rho([0,\infty), M_{1+|v|})$ to f for some $\rho > 0$ and $f \in C^1([0,\infty), M_{1+|v|})$ is the unique solution of (11). Furthermore $f_t \in M_{(1+|v|)^2}$ for all $t \in [0,\infty)$.*

By $M_{1+|v|}$ and $M_{(1+|v|)^2}$ we mean the set of Radon measures on $\mathbb{R}^d$ with first and second moments, C_ρ denotes the continuous functions which grow not faster than $e^{\rho t}$. The proof of Lemma 5 together with a precise definition of the function spaces can be found in Section 4.

Now we have to translate this idea into the context of deterministic many-body dynamics. To limit the complexity of the notation we will from now on assume that everything except the constants depends on a without displaying the dependency. For every realization of the N-body evolution the random variable $\beta(i,t) \in \{0,1\}$, which encodes the scattering state of particle $i \in \{1 \dots N\}$ at time $t \in [0,\infty)$, satisfies the implicit relation (8). The computation of β can be simplified by introducing a hierarchy of artificial evolutions indexed by $k \in \mathbb{N}$. We assume that the initial values of the particles at all levels are identical. The particles at level $k = 1$ are simply transported and do not interact with anything. The particles at level $k > 1$ interact only with the particles at level $k-1$, but not with each other. For each $k \in \mathbb{N}$ and $i \in \{1, \dots, N\}$ the scattering state $\beta_k(i,t) \in \{0,1\}$ is defined in the following way

$$\beta_k(i,t) = \begin{cases} 1 & \text{if } \operatorname{dist}(z_i, z_{i'}, s) \geq a\beta_{k-1}(i', s) \text{ for all } s \in [0,t), i' \neq i, \\ 0 & \text{else,} \end{cases} \tag{16}$$

$$\beta_1(i) \equiv 1, \tag{17}$$

with dist as in (9).

Remark 6. *While the determination of the collision-state $\beta(i,t)$ is a complicated problem, the state $\beta_k(i,t)$ emerges via a very simple calculation from $\beta_{k-1}(\cdot,t)$.*

Lemma 7. *For all realizations of the processes of the initial conditions $(u_0, v_0) \in \bigcup_{N=0}^{\infty} (\mathbb{T}^d \times \mathbb{R}^d)^N$ both $\beta_k(i,t)$ and $\beta(i,t)$ are well defined and*

$$\lim_{k\to\infty} \beta_k(i,t) = \beta(i,t) \tag{18}$$

pointwise in i and uniformly in t.

Proof. See section 4. □

2.2. **The concept of marked trees.** The translation of the N-body evolution into scattering states β is greatly facilitated by the concept of trees. In the collision tree with root (u,v) we will collect information of collisions and potential collisions up to time t for a particle with initial data (u,v).

As an example assume that $N = 4$ and consider the scenario in Fig. 1 where the letters A, B, C, D are the labels of the four particles, the empty circles are the initial positions and the arrows are the initial velocities. Consequently the arrow-tips indicate the positions of the particles at time $t = 1$.

To determine whether a certain particle has been scattered before time $t = 1$ it suffices to analyze the associated collision tree which is constructed as follows: The particle of interest is the root with initial data (u,v). The particles which are potentially scattered by the root are added as nodes, i.e. a particle with initial data (u',v') is added, if $\operatorname{dist}((u,v),(u',v'),s) \leq a$ for some $s \in [0,t]$. This procedure is recursively applied to every node but we consider only potential scattering events which are upstream, i.e. before the event which is responsible for adding the node. The four collision trees associated to the scenario in Fig. 1 are shown in Fig. 2. The extraction of the collision trees amounts to a significant reduction of the complexity of the problem. In general, the number of potential scattering events (bullets) grows with n. But thanks to the Boltzmann-Grad-scaling (2) the number of nodes in the individual trees is a random number related to a

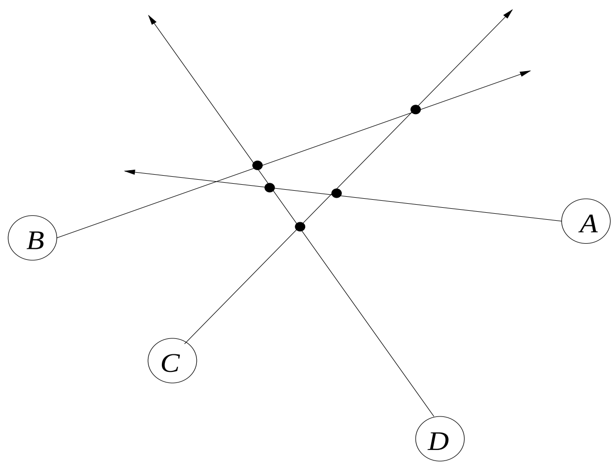

FIGURE 1. Initial positions and velocities of four particles. The bullets indicate the positions where the particles are potentially scattered. There is no bullet at the crossing of A and B as the particles would pass this point at different times. Given the high number of potential collisions, the shown configuration is not very likely and consequentially the collision trees are quite complex.

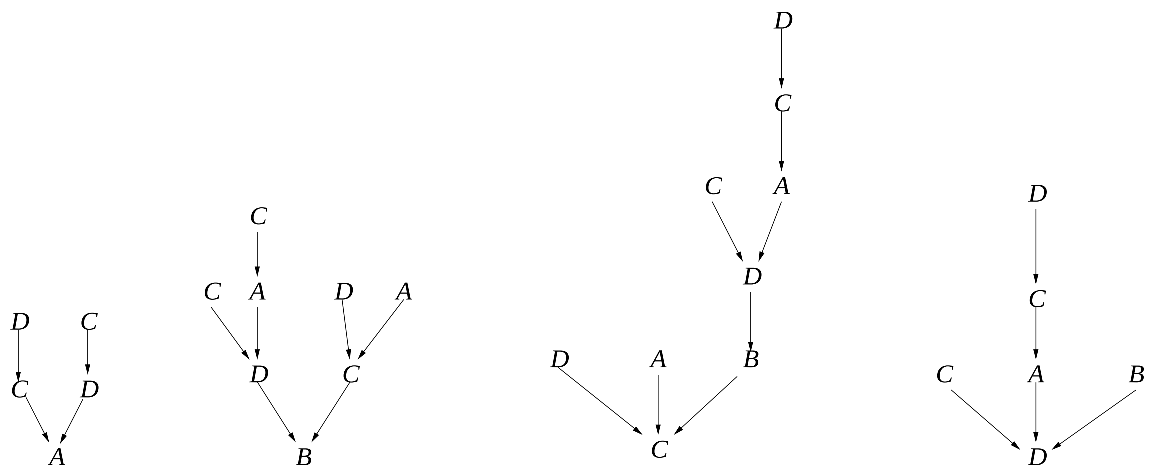

FIGURE 2. Collision trees of the four particles with initial positions and collision structure given in Fig. 1. At time $t = 1$ particles C and D have been scattered, particles A and B have not. The particle of interest is at the root. On the next level particles appear that (potentially) scatter the root particle. Particles are on the third level, if they (potentially) scatter particles on level two in the time of interest (until their collision with the root particle). This is iterated recursively. Note that the labels of the particles which generate the potential scattering events are only included in the picture in order to illustrate the translation of Fig. 1 into collision trees.

Poisson process with a distribution which is asymptotically independent of n and grows exponentially with t, see Lemma 13. In a physical interpretation this implies a constant "mean free path" in the scaling.
We convert now the example into a general concept.

Definition 8. *Let $\mathbb{N} = \{1, 2, \ldots\}$. The height of a node (or multi-index) $l \in \mathbb{N}^i$ is defined by $|l| := i$, the parent node of $l \in \mathbb{N}^i$ is $\bar{l} = (l_1, \ldots, l_{i-1})$. Let $\mathcal{F} = \cup_{i=1}^{\infty} \mathbb{N}^i$ be the set of multi-indices. We say that $m \subset \mathcal{F}$ is a tree skeleton with root ($m \in \mathcal{T}$), if*

(1) $\#m < \infty$,
(2) $m \cap \mathbb{N} = \{1\}$,
(3) $\bar{l} \in m$ *for all* $l \in m \setminus \{1\}$,
(4) $l - 1 \in m$ *for all* $l \in m$ *such that* $l \neq (*, \ldots, *, 1)$,

where $l - 1 = l - (0, \ldots, 0, 1)$. We say that a tree m has at most height k ($m \in \mathcal{T}_k$) if $m \cap \mathbb{N}^{k+1} = \emptyset$.
Let $Y = \{(u, v, s, \nu) \in \mathbb{T}^d \times \mathbb{R}^d \times [0, \infty) \times S^{d-1}\}$ be the space of initial values and collision parameters. The set of marked trees is given by

$$\mathcal{MT} = \Bigg\{ (m, \phi) \,\Bigg|\, m \in \mathcal{T}, \phi : m \to Y \text{ with the property } s_l \in [s_{l-1}, s_{\bar{l}}] \text{ and } \nu_l = \tfrac{1}{a}(u_{\bar{l}} - u_l + s_l(v_{\bar{l}} - v_l)) \text{ for all } l \in m \setminus \{1\} \Bigg\},$$

where $s_{(*,\ldots *,0)} = 0$. $\mathcal{MT}_k$ is obtained if $\mathcal{T}$ is replaced with $\mathcal{T}_k$. For each skeleton $m \in \mathcal{T}$ we define the set of marked trees with skeleton m

$$\mathcal{E}(m) = \{(\tilde{m}, \phi) \in \mathcal{MT} \mid \tilde{m} = m\}. \tag{19}$$

The assumption $s_l \in [s_{l-1}, s_{\bar{l}}]$ implies that for all nontrivial permutations $\pi \in S_{\#m} \setminus \mathrm{Id}$ (the set S_N contains the permutations of N symbols) and all trees $\Phi = (m, \phi) \in \mathcal{MT}$ the permuted tree $\Phi^\pi = (m, \phi^\pi)$ with $\phi_l^\pi = \phi_{\pi(l)}$ is not a tree in the sense of Definition 8.
The value ν_1 has no relevance. To circumvent this problem we fix a point $\nu^* \in (S^{d-1})$, define

$$\mathcal{MT}^* = \{\Phi \in \mathcal{MT} \mid \nu_1 = \nu_1^*\}.$$

We will in future denote $\mathcal{MT}^*$ by $\mathcal{MT}$. As an example consider the tree with A at its root in Fig. 2. The initial conditions are denoted by u_A, v_A, u_B, v_B etc. Then the marked tree is given by

$$\begin{aligned} m =& \{1, (1,1), (1,2), (1,1,1), (1,2,1)\} \\ \phi =& \{(1, (u_A, v_A, t, \nu^*)), ((1,1), (u_C, v_C, s_{11}, \nu_{11})), ((1,2), (u_D, v_D, s_{12}, \nu_{12})), \\ & ((1,1,1), (u_D, v_D, s_{111}, \nu_{12})), ((1,2,1), (u_C, v_C, s_{121}, \nu_{121}))\} \end{aligned} \tag{20}$$

where

$$\begin{aligned} s_{11} =& \min\{s \in [0,t] \mid \mathrm{dist}((u_A, v_A), (u_C, v_C), s) = a\} & \nu_{11} &= \tfrac{1}{a}(u_A - u_C + s_{11}(v_A - v_C)) \\ s_{12} =& \min\{s \in [0,t] \mid \mathrm{dist}((u_A, v_A), (u_D, v_D), s) = a\} & \nu_{12} &= \tfrac{1}{a}(u_A - u_D + s_{12}(v_A - v_D)) \\ s_{111} = s_{121} =& \min\{s \in [0,t] \mid \mathrm{dist}((u_C, v_C), (u_D, v_D), s) = a\} & \nu_{111} &= -\nu_{121} \\ & & &= \tfrac{1}{a}(u_D - u_C + s_{111}(v_D - v_C)) \end{aligned}$$

and dist is defined in eq. (9).
It is clear from the definition that for each tree $m \in \mathcal{T}$ there exists a function $r : m \to \mathbb{N} \cup \{0\}$ which counts the number of direct successors, i.e. for $l \in m$

$$r_l = \#\{l' \in m \mid \bar{l'} = l\}. \tag{21}$$

Remark 9. *Graph theoretical descriptions of collisions in a hard-sphere gas can lead to many different graphs, which are not necessarily trees. The advantage of our definition is that this graph will always be a tree. Particles might appear several times in a tree, as* *in Fig. 2. This will not destroy the tree structure, as these are due to different collision events. Multiple collisions, which are well-defined in our setting, can lead to identical branches within the tree, but the definition of $\mathcal{T}$ will discriminate between these and the graph of collisions is still a tree.*

Important information about the collisions of the root particle are encoded in the tree. In particular, a scattering state in $\{0,1\}$ is given by the tree. The scattering state of each node $l \in m$, which we also denote by $\beta : m \to \{0,1\}$, assigns to each node l the label 1 if it is unscattered by particles in the tree at time s_l and 0 if it was scattered before s_l. It is important to note that the scattering states of all particles described by nodes in the tree depend only on tree structure m, i.e. the scattering state is independent of the collision data ϕ, furthermore the scattering information relevant in the graph is completely determined by the state of the nodes on the higher levels: All leaves (nodes with no further successors/children $(r_l = 0)$) are assigned 1, as there are no collision events before s_l. Other nodes are assigned 0, if there exists at least one cild (l' such that $\bar{l'} = l$) with scattering state 1, i.e. there is real collision before s_l. The label 1 is assigned if all children have scattering state 0. Thus we define the scattering state $\beta : m \to \{0,1\}$ as follows.

$$\beta_l = \prod_{l' \in m, \bar{l'} = l} (1 - \beta_{l'}). \tag{22}$$

This definition rephrases the original definition of the scattering state in (16), adapting it to the tree structure. Here we drop the dependence on time as it is fixed for a tree and particles are replaced by nodes of a tree. In light of Lemma 15 below, we do not distinguish between the two notions.

We will construct now two families of probability measures $P_{t,k}, \hat{P}_{t,k} \in PM(\mathcal{MT}_k)$. The empirical distributions $\hat{P}_{t,k}$ describe the deterministic many-body dynamics with random initial data and will be constructed recursively in Section 2.4. The idealized distribution $P_{t,k}$ corresponds to the idealized statistical behavior as predicted by the Boltzmann equation (3). It is given by an explicit formula (24). The link between $P_{t,k}$ and $\hat{P}_{t,k}$ is provided by the set of good trees $\mathcal{G}(a) \subset \mathcal{MT}$ (Definition 18) which has the properties that restriction of $\hat{P}_{t,k}$ on $\mathcal{G}(a) \cap \mathcal{MT}$ converges to $P_{t,k}$ and $P_{t,k}(\mathcal{G}(a))$ goes to 1 as a tends to 0 (Proposition 22).

This is the crucial step which eventually yields the justification of the idealized theory. In other words, the main task consists in analyzing the idealized measure $P_{t,k}$, the empirical distribution $\hat{P}_{t,k}$ enters only when we prove that $P_{t,k}$ is consistent with $\hat{P}_{t,k}$.

2.3. **The idealized distribution $P_{t,k}$.** We construct now the idealized distribution of trees $P_{t,k} \in PM(\mathcal{MT})$. Viewing $\mathcal{E}(m) \subset (\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^+) \times \left(\mathbb{R}^d \times S^{d-1} \times \mathbb{R}^+\right)^{\#m-1}$ as a manifold, Borel sets $\Omega_m \subset \mathcal{E}(m)$ are well-defined. A set $\Omega \subset \mathcal{MT}$ is Borel, if

$$\Omega = \bigcup_{m \in \mathcal{T}} \Omega_m, \quad \Omega_m \subset \mathcal{E}(m) \text{ Borel.} \tag{23}$$

Let $\Omega \subset \mathcal{MT}$ be a Borel set and $t \in [0,\infty)$. The idealized probability that the observed tree is in Ω is given by

$$P_{t,k}(\Omega) = \sum_{m \in \mathcal{T}_k} \int_{\Omega \cap \mathcal{E}(m)} e^{-\sum_{j<k} \Gamma_j(\Phi)} \,\mathrm{d}\lambda^m(\phi) \tag{24}$$

where $\mathcal{E}(m)$ was defined in (19)

$$\Gamma_j(\Phi) = \sum_{l\in m, |l|=j} \gamma_l(\Phi), \tag{25}$$

$$\gamma_l(\Phi) = \int_0^{s_l} L[f_0](v_l)\,\mathrm{d}s' = s_l\, L[f_0](v_l) \geq 0 \text{ is the integrated collision rate of the particle at node } l, \tag{26}$$

$$\lambda^m(\phi) = \mathbf{1}_{\mathbb{T}^d}(u_1) \otimes f_0(v_1) \otimes \delta(s_1 - t) \otimes \prod_{l\in m\setminus\{1\}} \left[((v_l - v_{\bar{l}}) \cdot \nu_l)_+ \, \chi_{[s_{l-1}, s_{\bar{l}}]}(s_l)\,\mathrm{d}f_0(v_l)\,\mathrm{d}\nu_l\,\mathrm{d}s_l\,\right]. \tag{27}$$

Remark 10. (1) *Note that the positions u_l are completely determined by (u_1, v_1) and $(v_l, s_l, \nu_l)_{l\in m\setminus\{1\}}$. Since we have assumed that (ν_1) is fixed, the value of $P_{t,k}(\Omega)$ is well-defined.*

(2) *In (26) we assign to each node l a particle, this map might not be injective, e.g. in the example (20) both $(1, 1)$ and $(1, 2, 1)$ would refer to the same particle but different times.*

(3) *It is noteworthy that the measures $P_{t,k}$ depend on time only via the parameter t. In other words, time plays the role of a parameter which propagates through the tree and qualifies the local branching structure.*

(4) *For some event $\Omega \subset \mathcal{MT}_k$ the probability $P_{t,k'}(\Omega)$ is independent of k' if $k' > k$. Equivalently, $P_{t,k_1}(\Omega \cap \mathcal{E}(m)) = P_{t,k_2}(\Omega \cap \mathcal{E}(m))$, if the height of m is strictly smaller than $\min\{k_1, k_2\}$.*

(5) *Clearly $P_{t,1}$ is a probability measure. It follows from Lemma 12 below with $x(m) = 1$ that $P_{t,k}$ is a probability measure for all (t, k).*

We can simplify the measure $P_{t,k}$ by integrating over the collision parameters $\nu_l \in S^{d-1}$, $l \in m$. Let $\hat{Y} = \mathbb{R}^d \times [0, \infty)$ be the reduced set of collision data and $\widehat{\mathcal{MT}}$ the corresponding marked trees. For every $\Omega \subset \widehat{\mathcal{MT}}$ we find that when still denoting the collision data as ϕ

$$\bar{P}_{t,k}(\Omega) = \sum_{m\in\mathcal{T}_k} \int_{\Omega\cap\mathcal{E}(m)} e^{-\sum_{j<k}\Gamma_j(\Phi)}\mathrm{d}f_0(v_1) \otimes \delta(s_1 - t) \otimes \prod_{l\in m\setminus\{1\}} \mathrm{d}\bar{\lambda}_l(\phi) \tag{28}$$

with

$$\mathrm{d}\bar{\lambda}_l(\phi) = \kappa_d\, |v_l - v_{\bar{l}}|\, \chi_{[s_{l-1}, s_{\bar{l}}]}(s_l)\,\mathrm{d}f_0(v_l)\,\mathrm{d}s_l. \tag{29}$$

The measures $P_{t,k}$ have the remarkable property that the expectation of certain random variables can be computed efficiently.

Definition 11. *A random variable $x : \mathcal{T} \to \mathbb{R}$ is said to be recursive if there exists a family of functions $h_b : \mathbb{R}^b \to \mathbb{R}$, $b \in \mathbb{N}$, which are invariant under permutations of the b components in $\mathbb{R}^b$, such that for all $m \in \mathcal{T}$ with $b = r_1$ as defined in (21) the equation*

$$x(m) = h_{r_1}(x(m_1), \ldots, x(m_{r_1}))$$

holds, where

$$m_j = \{(1, l_3, \ldots, l_{|l|}) \mid l \in m \text{ such that } l_2 = j\} \in \mathcal{T}$$

is the j-th subtree of m.

Examples of recursive random variables which are relevant for our purposes are

$$x^{\#}(m) = \#m \text{ (number of nodes)},$$

$$x^{\beta}(m) = \beta_1(m) \text{ (scattering state of the root)}.$$

It is easy to see that if $m \in \mathcal{T}$

$$x^{\#}(m) = 1 + \sum_{j=1}^{r_1} x^{\#}(m_j), \tag{30}$$

$$x^{\beta}(m) = \prod_{j=1}^{r_1}(1 - x^{\beta}(m_j)) \text{ with the convention } \prod_{j=1}^{0}(1 - x^{\beta}(m_j)) = 1, \tag{31}$$

which both depend on the tree structure m alone. Hence the functions h_b are given by

$$h_b^{\#}(x_1, \ldots, x_b) = 1 + \sum_{j=1}^{b} x_j,$$

$$h_b^{\beta}(x_1, \ldots, x_b) = \prod_{j=1}^{b}(1 - x_j)$$

which are clearly invariant under permutations of $x_1, \ldots, x_b$. The expectation of recursive random variables with respect to the probability measure $P_{t,k}$ can be computed with a simple recurrence relation.

Lemma 12. *Let x be a recursive random variable with recurrence functions h_b. Then*

(32)

$$\begin{aligned}
&\int \mathrm{d}\bar{P}_{t,k}(\Phi)\, x(m) \\
&= \int \mathrm{d}f_0(v)\, e^{-\Gamma_1} \sum_{r=0}^{\infty} \int_0^t \mathrm{d}s_1 \int \mathrm{d}\bar{P}_{s_1,k-1}(\Phi_1)\, \kappa_d |v - v_1| \int_{s_1}^{t} \mathrm{d}s_2 \int \mathrm{d}\bar{P}_{s_2,k-1}(\Phi_2)\, \kappa_d |v - v_2| \\
&\quad \ldots \int_{s_{r-1}}^{t} \mathrm{d}s_r \int \mathrm{d}\bar{P}_{s_r,k-1}(\Phi_r)\, \kappa_d |v - v_r|\, h_r(x(m_1), \ldots, x(m_r)) \\
&= \int \mathrm{d}f_0(v)\, e^{-\Gamma_1} \sum_{r=0}^{\infty} \frac{1}{r!} \int_0^t \mathrm{d}s_1 \int \mathrm{d}\bar{P}_{s_1,k-1}(\Phi_1)\, \kappa_d |v - v_1| \int_{0}^{t} \mathrm{d}s_2 \int \mathrm{d}\bar{P}_{s_2,k-1}(\Phi_2)\, \kappa_d |v - v_2| \\
&\quad \ldots \int_{0}^{t} \mathrm{d}s_r \int \mathrm{d}\bar{P}_{s_r,k-1}(\Phi_r)\, \kappa_d |v - v_r|\, h_r(x(m_1), \ldots, x(m_r))
\end{aligned}$$

where for $r = 0$ we assign 1 to the empty product, v is the velocity of the root particle, v_j denotes the velocity of the root particle of the subtree $\Phi_r = (m_r, \phi)$ and $\Gamma_1 = \kappa_d \int \mathrm{d}f_0(v')\, |v - v'| t$.

Proof. As $x(m)$ does not depend on the collision parameter ν_l, we can restrict our attention to $\bar{P}_{t,k}$ and $\widehat{\mathcal{MT}}$. For each $\Phi \in \widehat{\mathcal{MT}}$ ($\hat{Y} = \mathbb{R}^d \times [0, \infty)$) we define nonnegative Radon measures $\bar{\lambda}_l \in M_+(\mathbb{R}^d \times [0, \infty))$ as in (29)

$$\mathrm{d}\bar{\lambda}_l(v, s) = \mathrm{d}f_0(v)\, \kappa_d |v_{\bar{l}} - v|\, \chi_{[s_{l-1}, s_{\bar{l}}]}(s) \mathrm{d}s.$$

Let now $m \in \mathcal{T}$. The definition of $\bar{P}_{t,k}$ in (28) yields

$$\int_{\mathcal{E}(m)} \mathrm{d}\bar{P}_{t,k}(\Phi)\, x(m) = \int_{\mathcal{E}(m)} e^{-\sum_{j<k} \Gamma_j(\Phi)}\, \mathrm{d}f_0(v_1) \prod_{i=1}^{r_1} \left[\mathrm{d}\bar{\lambda}_{1i}(\phi_{1i}) \prod_{\substack{l \in m \setminus (\{1\} \cup \mathbb{N}^2) \\ l_2 = i}} \mathrm{d}\bar{\lambda}_l(\phi_l) \right] x(m),$$

We use now the assumption that x is recursive and find

$$\begin{aligned}&\int_{\mathcal{E}(m)} \mathrm{d}\bar{P}_{t,k}(\Phi)\, x(m)\\ &= \int \mathrm{d}f_0(v)\, e^{-\Gamma_1} \prod_{i=1}^{r_1} \left[\int_{\mathcal{E}(m_i)} e^{-\sum_{j<k} \Gamma_j^{(i)}(\Phi)} \prod_{l\in m_i\setminus\{1\}} \mathrm{d}\bar{\lambda}_l(\phi_l) \right] h_{r_1}(x(m_1),\dots,x(m_{r_1})),\end{aligned}$$

where $\Gamma_j^{(i)}(\Phi) = \sum\limits_{l\in m,|l|=j,l_2=i} \gamma_l(\phi)$. A simple rearrangement yields that

$$\begin{aligned}\sum_{m\in\mathcal{T}} \int_{\mathcal{E}(m)} \mathrm{d}\bar{P}_{t,k}(\Phi)\, x(m) = \int \mathrm{d}f_0(v)\, e^{-\Gamma_1} \sum_{r=0}^{\infty} \int_0^t \mathrm{d}s_1 \int \mathrm{d}\bar{P}_{s_1,k-1}(\Phi_1)\, \kappa_d |v-v_1|\\ \dots \int_{s_{r-1}}^t \mathrm{d}s_r \int \mathrm{d}\bar{P}_{s_r,k-1}(\Phi_r)\, \kappa_d|v-v_r|\, h_r(x(m_1),\dots,x(m_r))\,,\end{aligned}$$

where v_j denotes the velocity of the root particle of the subtree $\Phi_r = (m_r,\phi)$. This demonstrates the first part of (32), to show the second part we observe that

$$\begin{aligned}&\{(s_1,\dots,s_r)\in[0,t]^r \mid s_j\neq s_i \text{ for } i\neq j\}\\ &= \bigcup_{\pi\in S_r} \{(s_1,\dots,s_r)\in[0,t]^r \mid s_{\pi(1)}<s_{\pi(2)}<\dots<s_{\pi(r)}\},\end{aligned}$$

where S_r denotes the symmetric group on r elements, such that the union is disjoint. As the set, where $s_j = s_i$ for some $i\neq j$, is of measure zero with respect to Lebesgue measure on $[0,t]^r$, we obtain

$$\int_{[0,t]^r} g(s_1,\dots,s_r)\,\mathrm{d}s_1\dots\mathrm{d}s_r = \sum_{\pi\in S_r} \int_{0\le s_{\pi(1)}<s_{\pi(2)}<\dots<s_{\pi(r)}\le t} g(s_1,\dots,s_r)\,\mathrm{d}s_1\dots\mathrm{d}s_r$$

for any $g\in L^1([0,t]^r)$. Now we define

$$g(s_1,\dots,s_r) = \int \mathrm{d}\bar{P}_{s_1,k-1}(\Phi_1)\,\kappa_d|v-v_1|\dots\int \mathrm{d}\bar{P}_{s_r,k-1}(\Phi_r)\,\kappa_d|v-v_r|\,h_r(x(m_1),\dots x(m_r)).$$

We observe that

$$\begin{aligned}&\bar{P}_{s_1,k-1}(\Phi_1)\,\kappa_d|v-v_1|\dots\bar{P}_{s_r,k-1}(\Phi_r)\,\kappa_d|v-v_r|\\ =&\bar{P}_{s_{\pi(1)},k-1}(\Phi_{\pi(1)})\,\kappa_d|v-v_{\pi(1)}|\dots\bar{P}_{s_{\pi(r)},k-1}(\Phi_{\pi(r)})\,\kappa_d|v-v_{\pi(r)}|\end{aligned} \tag{33}$$

for all permutations $\pi\in S_r$. Next using (33) and the invariance of h under permutations, we obtain

$$\begin{aligned}&\int_{0\le s_1<s_2<\dots<s_r\le t} \int \mathrm{d}\bar{P}_{s_1,k-1}(\Phi_1)\,\kappa_d|v-v_1|\\ &\dots\int \mathrm{d}\bar{P}_{s_r,k-1}(\Phi_r)\,\kappa_d|v-v_r| h_r(x(m_1),\dots x(m_r))\,\mathrm{d}s_1\dots\mathrm{d}s_r\\ =&\int_{0\le s_{\pi(1)}<s_{\pi(2)}<\dots<s_{\pi(r)}\le t} \int \mathrm{d}\bar{P}_{s_{\pi(1)},k-1}(\Phi_{\pi(1)})\,\kappa_d|v-v_{\pi(1)}|\\ &\dots\int \mathrm{d}\bar{P}_{s_{\pi(r)},k-1}(\Phi_{\pi(r)})\,\kappa_d|v-v_r| h_r(x(m_{\pi(1)}),\dots x(m_{\pi(r)}))\,\mathrm{d}s_1\dots\mathrm{d}s_r.\end{aligned}$$

As there are $r!$ different permutations in S_r we finally have

$$
\begin{aligned}
&\int_{0\le s_1<s_2<\ldots<s_r\le t}\int \mathrm{d}\bar P_{s_1,k-1}(\Phi_1)\,\kappa_d|v-v_1|\\
&\quad\ldots\int \mathrm{d}\bar P_{s_r,k-1}(\Phi_r)\,\kappa_d|v-v_r|\,h_r(x(m_1),\ldots x(m_r))\,\mathrm{d}s_1\ldots\mathrm{d}s_r\\
=&\frac{1}{r!}\int_{[0,t]^r}\int \mathrm{d}\bar P_{s_1,k-1}(\Phi_1)\,\kappa_d|v-v_1|\\
&\quad\ldots\int \mathrm{d}\bar P_{s_r,k-1}(\Phi_r)\,\kappa_d|v-v_r|\,h_r(x(m_1),\ldots x(m_r))\,\mathrm{d}s_1\ldots\mathrm{d}s_r.
\end{aligned}
$$

Summing over r and m completes the proof of (32). □

As an application of Lemma 12 we obtain an explicit bound on the expected number of nodes in trees.

Lemma 13. *For a tree $m\in\mathcal{T}$ the number of non-root nodes is given by $X(m)=\#m-1$. The expected value of X with respect to measure $P_{t,k}$ satisfies the estimate uniformly in k*

$$
\mathbb{E}(X)\le K_{\mathrm{ini}}\exp(\kappa_d K_{\mathrm{ini}}t), \tag{34}
$$

with $K_{\mathrm{ini}}=\int_{\mathbb{R}^d}\mathrm{d}f_0(v)\,(1+|v|)^2$ as in (5).

Proof. Let $F_{t,k}(v)=\mathbb{E}(X\mid v_1=v,m\in\mathcal{T}_k)$ be the conditional expectation of X if we know that the velocity of the root is v and that the tree is in $\mathcal{T}_k$. Clearly $\mathbb{E}(X)\le\sup_{k\in\mathbb{N}}\int_{\mathbb{R}^d}\mathrm{d}f_0(v)\,F_{t,k}(v)$. Now we use the self-similarity relation (32) with $x(m)=X(m)$ and $h_r(X(m_1),\ldots,X(m_r))=r+\sum_{i=1}^r X(m_i)$. The velocity of the root particle of m_i is denoted by $v_1^{(i)}$ and we let as in (25)

$$
\Gamma_1(v_1)=\gamma_1(v_1)=L[f_0](v_1)t=\kappa_d t\int_{\mathbb{R}^d}\mathrm{d}f_0(v')\,|v_1-v'|. \tag{35}
$$

Then

$$
\begin{aligned}
&F_{t,k}(v_1)\\
=&e^{-\Gamma_1}\sum_{r=1}^\infty\tfrac{1}{r!}\int_0^t\mathrm{d}s_1\int\mathrm{d}\bar P_{s_1,k-1}(m_1,\phi_1)\,\kappa_d|v_1-v_1^{(1)}|\\
&\ldots\int_0^t\mathrm{d}s_r\int\mathrm{d}\bar P_{s_r,k-1}(m_r,\phi_r)\,\kappa_d|v_1-v_1^{(r)}|\left(r+\sum_{i=1}^r X(m_i)\right)\\
=&e^{-\Gamma_1}\sum_{r=1}^\infty\left(r\frac{\Gamma_1^r}{r!}+\frac{\Gamma_1^{r-1}}{r!}\sum_{i=1}^r\int_0^t\mathrm{d}s_i\,\kappa_d\int_{\mathbb{R}^d}\mathrm{d}f_0(v_1^{(i)})\,|v_1-v_1^{(i)}|\,F_{s_i,k-1}(v_1^{(i)})\right)\\
=&\Gamma_1+\int_0^t\mathrm{d}s\,\kappa_d\int_{\mathbb{R}^d}\mathrm{d}f_0(v')\,|v_1-v'|\,F_{s,k-1}(v'),
\end{aligned}
$$

where we used the product structure of the integrals and (35) to obtain e.g.

$$
\begin{aligned}
&\int_0^t\mathrm{d}s_j\int\mathrm{d}\bar P_{s_j,k-1}(m_j)\,\kappa_d|v_1-v_1^{(j)}|X(m_j)\prod_{i=1,\ldots,r;i\neq j}\int_0^t\mathrm{d}s_i\int\mathrm{d}\bar P_{s_i,k-1}(m_i)\,\kappa_d|v_1-v_1^{(i)}|\\
=&\Gamma_1^{r-1}\int_0^t\mathrm{d}s_j\,\kappa_d\int_{\mathbb{R}^d}\mathrm{d}f_0(v_1^{(j)})\,|v_1-v_1^{(i)}|\,F_{s_j,k-1}(v_1^{(j)}).
\end{aligned}
$$

We define now the norm $\|F\|_1 := \sup_{v\in\mathbb{R}^d} \frac{F(v)}{1+|v|}$ and the integral operator A_{f_0} by

$$(A_{f_0}F)(v) = \kappa_d \int_{\mathbb{R}^d} \mathrm{d}f_0(v')\,|v-v'|F(v'),$$

so that

$$F_{t,k} = \Gamma_1 + \int_0^t \mathrm{d}s\, A_{f_0} F_{s,k-1}. \tag{36}$$

We find the estimates

$$\|A_{f_0}F\|_1 \le \sup_v \frac{\kappa_d\|F\|_1}{1+|v|} \int_{\mathbb{R}^d} \mathrm{d}f_0(v')\,|v-v'|\,(1+|v'|) \le \kappa_d K_{\mathrm{ini}}\|F\|_1,$$

and

$$\|\Gamma_1\|_1 = t\sup_v \frac{\kappa_d \int_{\mathbb{R}^d} \mathrm{d}f_0(v')\,|v-v'|}{1+|v|} \le t\kappa_d \int_{\mathbb{R}^d} \mathrm{d}f_0(v')\,(1+|v'|) \le t\kappa_d K_{\mathrm{ini}}.$$

Furthermore $F_{t,k}(v)$ is monotone in k, as $P_{t,k}$ assigns the probability of trees of height greater than $k+1$ to trees of height k, reducing the number of expected nodes. Hence eq. (36) implies that

$$\|F_{t,k}\|_1 \le \kappa_d K_{\mathrm{ini}} \left(t + \int_0^t \mathrm{d}s\, \|F_{s,k}\|_1 \right).$$

Gronwall's inequality together with the previous estimate implies that

$$\|F_{t,k}\|_1 \le e^{\kappa_d K_{\mathrm{ini}} t},$$

where we used that $F_{0,k} = 0$. Since

$$\mathbb{E}_k(X) = \int_{\mathbb{R}^d} \mathrm{d}f_0(v)\, F_{t,k}(v) \le \|F_{t,k}\|_1 \int_{\mathbb{R}^d} \mathrm{d}f_0(v)\,(1+|v|) \le K_{\mathrm{ini}} e^{\kappa_d K_{\mathrm{ini}} t}$$

this implies (34) and the proof of the lemma is finished. □

We now turn our attention to the determination of the scattering state of the particle at the root of the tree. For a tree $m \in \mathcal{T}$ the scattering state $\beta : m \to \{0,1\}$ is defined recursively by (22). It is more convenient in our analysis than the ad-hoc definition, which required already some work to show existence, see the first part of Lemma 7. The important simplification is that scattering state in (22) only depends on the structure m but is independent of the data ϕ.
We define the single-particle density $g_{t,k}(\cdot) \in M_+(\mathbb{R}^d)$ via

$$\int_A \mathrm{d}g_{t,k}(v) = P_{t,k}(\beta_1 = 1 \text{ and } v_1 \in A),$$

for all $A \subset \mathbb{R}^d$ Borel. The density $g_{t,k}$ is closely related to the root marginal of $P_{t,k}$ and provides the link between the Boltzmann equation (11) and the idealized distribution of the trees $P_{t,k}$. Due to the simplicity of the distribution $P_{t,k}$ it is possible to characterize the root-marginal of $P_{t,k}$ explicitly.

Proposition 14. *Let $\sigma \in \{0,1\}$, $\Omega \subset \mathbb{R}^d$ Borel, $t \in [0,\infty)$ and $k \in \mathbb{N}\cup\{0\}$. Then the equation*

$$P_{t,k+1}(v_1 \in \Omega \text{ and } \beta_1 = \sigma_1) = \int_\Omega \left[(1-\sigma_1)\,(\mathrm{d}f_0(v) - \mathrm{d}f_{t,k}(v)) + \sigma_1\,\mathrm{d}f_{t,k}(v)\right] \tag{37}$$

holds, where $f_{t,k}$ is the solution of system (15).

This formula shows that in particular $g_{t,k} = f_{t,k-1}$.

Proof. The proposition is proven using induction over k, the case $k=0$ is just the definition. In the induction step it is demonstrated that $P_{t,k+1}$ satisfies formula (37) if $P_{t,k}$ does. Since the collision parameters ν are irrelevant we can integrate them out and work with the simplified version (28) of the measure $P_{t,k}$ instead of (24).
We define the set of scattering states of trees up to height 2 that are compatible with $\sigma \in \{0,1\}$,

$$\mathcal{A}(\sigma) = \left\{ (m,\sigma') \;\middle|\; m \in \mathcal{T}_2, \sigma' : m \to \{0,1\} \text{ such that } \prod_{l' \in m \cap \mathbb{N}^2} (1-\sigma'_l) = \sigma \right\}, \tag{38}$$

with the standard convention $\prod_{j=1}^{0} a_j = 1$ for empty products, i.e.

$$\begin{aligned}
\mathcal{A}(0) &= \left\{ (m,\sigma') \;\middle|\; m \in \mathcal{T}_2, \sigma' : m \to \{0,1\} \text{ such that } \sigma'_l = 1 \text{ for some } l \in m \cap \mathbb{N}^2 \right\}, \\
\mathcal{A}(1) &= \left\{ (m,\sigma') \;\middle|\; m \in \mathcal{T}_2, \sigma' : m \to \{0,1\} \text{ such that } \sigma'_l = 0 \text{ for all } l \in m \cap \mathbb{N}^2 \right\}.
\end{aligned}$$

The induction assumption and eq. (32) implies that

$$\begin{aligned}
&P_{t,k+1}(v_1 \in \Omega \text{ and } \beta_1 = \sigma) \\
&= \sum_{(m,\sigma')\in\mathcal{A}(\sigma)} \int_{v_1\in\Omega} \Bigg(\tfrac{e^{-\Gamma_1}}{r_1!} \,\mathrm{d}f_0(v_1) \prod_{l'\in m\cap\mathbb{N}^2} \Bigg[(1-\sigma'_{l'}) \int_0^{s_1} \mathrm{d}s \int_{v'\in\mathbb{R}^d} \kappa_d |v_1 - v'| \left(\mathrm{d}f_0(v') - \mathrm{d}f_{s,k-1}(v')\right) \\
&\qquad\qquad + \sigma'_{l'} \int_0^{s_1} \mathrm{d}s \int_{v'\in\mathbb{R}^d} \mathrm{d}f_{s,k-1}(v')\, \kappa_d |v_1 - v'| \Bigg] \Bigg) \\
&= \int_{v_1\in\Omega} \sum_{(m,\sigma')\in\mathcal{A}(\sigma)} \mathrm{d}f_0(v_1)\, I_k(m,\sigma',v_1),
\end{aligned}$$

where

$$\begin{aligned}
&I_k(m,\sigma',v_1) \\
&= \tfrac{e^{-\Gamma_1}}{r_1!} \prod_{l'\in m\cap\mathbb{N}^2} \Bigg[(1-\sigma'_{l'}) \int_0^{t} \mathrm{d}s \int_{v'\in\mathbb{R}^d} \kappa_d |v_1 - v'| (\mathrm{d}f_0(v') - \mathrm{d}f_{s,k-1}(v')) \\
&\qquad + \sigma'_{l'} \int_0^{t_l} \mathrm{d}s \int_{v'\in\mathbb{R}^d} \mathrm{d}f_{s,k-1}(v')\, \kappa_d |v_l - v'| \Bigg] \\
&= \tfrac{e^{-\Gamma_1}}{r_1!} \prod_{l'\in m\cap\mathbb{N}^2} \Bigg[(1-\sigma'_{l'}) \left(\Gamma_1 - \int_0^t \mathrm{d}s\, L[f_{s,k-1}](v_1) \right) + \sigma'_{l'} \int_0^t \mathrm{d}s\, L[f_{s,k-1}](v_1) \Bigg].
\end{aligned}$$

We rewrite $P_{t,k+1}(\ldots)$ as follows:

$$P_{t,k+1}(v_1 \in \Omega \text{ and } \beta_1 = \sigma) = \int_{v_1\in\Omega} \mathrm{d}f_0(v_1) \left[(1-\sigma) J_k(0,v_1) + \sigma J_k(1,v_1) \right], \tag{39}$$

with $J_k(\sigma,v_1) = \sum_{(m,\sigma')\in\mathcal{A}(\sigma)} I_k(m,\sigma',v_1)$. By definition $\mathcal{A}(1)$ assigns to each skeleton $m \in \mathcal{T}_2$ a unique σ' which assumes the value 1 on the root and 0 on all nodes on the second level (this includes the special case $m' = (1) \in \mathcal{T}_2$, that has no nodes on the second level). This shows that

$$J_k(1,v_1) = \sum_{j=0}^{\infty} \frac{e^{-\gamma}}{j!} \left(\gamma - \int_0^s \mathrm{d}s'\, L[f_{s',k-1}](v_1) \right)^j = e^{-\int_0^s \mathrm{d}s'\, L[f_{s',k-1}](v_1)}, \tag{40}$$

with $\gamma = sL[f_0](v_1)$. As

$$\begin{aligned}\int_\Omega \mathrm{d}f_0(v) &= P_{t,k}(v_1 \in \Omega \text{ and } \beta_1 \in \{0,1\}) \\ &= P_{t,k}(v_1 \in \Omega \text{ and } \beta_1 = 0) + P_{t,k}(v_1 \in \Omega \text{ and } \beta_1 = 1) \\ &= \int_\Omega \mathrm{d}f_0(v_1)\, J_k(0,v_1) + \int_\Omega \mathrm{d}f_0(v_1)\, J_k(1,v_1),\end{aligned}$$

we have

(41)

$$\int_{v_1\in\Omega} \mathrm{d}f_0(v_1)\, J_k(0,v_1) = \int_{v_1\in\Omega} \mathrm{d}f_0(v_1)\,(1 - J_k(1,v_1)) = \int_{v_1\in\Omega} \mathrm{d}f_0(v_1)\left(1 - e^{-\int_0^s \mathrm{d}s'\, L[f_{s',k-1}]}\right).$$

Plugging the formulas (40) and (41) into eq. (39) yields that

$$\begin{aligned}&P_{t,k+1}(v_1 \in \Omega \text{ and } \beta_1 = \sigma_1) \\ &= \int_{v_1\in\Omega} \left[(1-\sigma_1)\,\mathrm{d}f_0(v_1)\,\left(1 - e^{-\int_0^{t_1} \mathrm{d}s\, L[f_{s,k-1}]}\right) + \sigma_1\, \mathrm{d}f_0(v_1)\, e^{-\int_0^{t_1} \mathrm{d}s\, L[f_{s,k-1}]}\right] \\ &\overset{(15)}{=} \int_{v_1\in\Omega} (1-\sigma_1)\,(\mathrm{d}f_0(v_1) - \mathrm{d}f_{t_1,k}(v_1)) + \sigma_1\, \mathrm{d}f_{t_1,k}(v_1)\end{aligned}$$

and formula (37) has been established. □

2.4. **The empirical distribution $\hat{P}_{t,k}$.** We return now to the hierarchy of many body evolutions described in Section 2.1. The initial values of the particles form a random set $\omega \subset \mathbb{T}^d \times \mathbb{R}^d$ and it is assumed that the law of ω is the Poisson point process with density $\mu = n(\mathbf{1}_{\mathbb{T}^d} \otimes f_0)$, where $\mathbf{1}_{\mathbb{T}^d} \otimes f_0 \in PM(\mathbb{T}^d \times \mathbb{R}^d)$. Hence, the size $N = \#\omega$ is Poissonian random variable with intensity n. As explained in Section 2.2, the family of probability measures $\hat{P}_{t,k} \in PM(\mathcal{MT})$ is the empirical distribution of the tree Φ which is generated by the many-body evolution and has a randomly chosen (tagged) particle as its root. The scattering state of the root gives the connection between (16) and (22).

Lemma 15. *Let $\Phi = (m,\phi) \in \mathcal{MT}_k$ and i^* is the index the root particle in (16) then $\beta_1 = \beta_k(i^*, s_1)$.*

Proof. See section 4. □

The method of sampling from this distribution consists in drawing a realization of ω according to the unconditioned Poisson point process, and an independent random variable $z \in \mathbb{T}^d \times \mathbb{R}^d$ with law $\mathbf{1}_{\mathbb{T}^d}(u) \otimes f_0(v)$ which is the initial value of the tagged particle. The trees generated by this procedure are denoted by $\Phi(t,k) = (m(t,k),\phi) \in \mathcal{MT}_k$, where $m(t,k) \in \mathcal{T}_k$ is the skeleton and $\phi : m(t,k) \to Y$ specifies the initial values, the collision times and the impact parameters. The measures $\hat{P}_{t,k}$ are the image measure of $\mathrm{Prob}_{\mathrm{tppp}}$ induced by the many-particle flows so that for each Borel set $\Omega \subset \mathcal{MT}$ in the sense of (23) we obtain

$$\hat{P}_{t,k}(\Omega) := \mathrm{Prob}_{\mathrm{tppp}}((m(t,k),\phi) \in \Omega). \tag{42}$$

By construction, for fixed ω the skeleton m is monotonously increasing in t and k, and for fixed $l \in m$ the data ϕ_l does not depend on t or k. This implies that the j-marginal of $\hat{P}_{t,k}$ (trees of height $j \le k$) is given by $\hat{P}_{t,j}$, i.e.

$$\hat{P}_{t,k}\left((m(t,k) \cap (\cup_{i=1}^j \mathbb{N}^i), (\phi_l)_{|l|\le j}) \in \Omega\right) = \hat{P}_{t,j}((m(t,j),(\phi_l)_{|l|\le j}) \in \Omega) \tag{43}$$

for all $\Omega \subset \mathcal{MT}_j$, $k \ge j$.

We will use formula (43) to construct an alternative characterization of $\hat{P}_{t,k}$ which reflects the iterative process that underlies the definition of $m(t,k)$. Using this alternative characterization one can easily establish total-variation bounds for $P_{t,k} - \hat{P}_{t,k}$. Since the time t is arbitrary but fixed we will often write $\hat{P}_k$ instead of $\hat{P}_{t,k}$.
Let $(m', \phi') \in \mathcal{MT}_{k-1}$ and let $\hat{P}_k(\,\cdot\,|(m',\phi')) \in PM(\mathcal{MT}_k)$ be the conditional distribution of $\hat{P}_k$ in the sense that

$$\hat{P}_k(\Omega\,|\,(m',\phi')) := \hat{P}_k\Big((m(k),\phi) \in \Omega \mid m \cap \mathbb{N}^j = m' \cap \mathbb{N}^j \text{ for all } j \in \{1 \dots k-1\} \\ \text{and } \phi_l = \phi'_l \text{ for all } l \in m \text{ such that } |l| < k\Big).$$

Formula (43), which characterizes the j-marginals of $\hat{P}_{t,k}$, yields the following recurrence relation for $\hat{P}_k$:

$$\hat{P}_k(\Omega) = \int_{\mathcal{MT}_{k-1}} \mathrm{d}\hat{P}_{k-1}(\Phi')\, \hat{P}_k(\Omega\,|\,\Phi'). \tag{44}$$

Repeating this step $k-1$ times we obtain the following iterative representation of $\hat{P}_k$:
(45)

$$\hat{P}_k(\Omega) = \int_{\mathcal{MT}} \mathrm{d}P_1(\Phi_1) \int_{\mathcal{MT}_2} \mathrm{d}\hat{P}_2(\Phi_2\,|\,\Phi_1) \dots \int_{\mathcal{MT}_{k-1}} \mathrm{d}\hat{P}_{k-1}(\Phi_{k-1}\,|\,\Phi_{k-2})\, \hat{P}_k(\Omega\,|\,\Phi_{k-1}),$$

where

$$P_1(z_1) = (\mathbf{1}_{\mathbb{T}^d} \otimes f_0)(z_1) \in PM(\mathbb{T}^d \times \mathbb{R}^d) \tag{46}$$

is the distribution of initial values.

Remark 16. *Equation (45) shows that* $\mathrm{d}\hat{f}^{(a)}_{t,k}(u,v) = \hat{P}_{t,k}\big(z(1,t) \in [u, u+\mathrm{d}u) \times [v, v+\mathrm{d}v) \text{ and } \beta^{(a)}(1,t) = 1\big)$ *is absolutely continuous with respect to* $\mathbf{1}_{\mathbb{T}^d} \otimes f_0$.

2.5. **Representation of $\hat{P}_k - P_k$.** Having constructed an iterative characterization of $\hat{P}_k$ we will now show that it is very similar to the idealized measure P_k in a precise way. The key is to identify the mechanisms by which the two probability distributions fail to be equal. In this part of the paper we will work with the phase-space representation of the trees: $z_l = (u_l, v_l) \in \mathbb{T}^d \times \mathbb{R}^d$.

Remark 17. *There are only two reasons why $\hat{P}_k$ fails to coincide with P_k in the limit $a \to 0$:*

1. *The cylinders which are covered by the paths of the particles might contain self-intersections due to the periodic boundary conditions: $v - v' \in R(t,a)$ with*

$$R(t,a) = \left\{ v \in \mathbb{R}^d \mid \min\{|s\,v - \xi| \mid s \in [0,t], \xi \in \mathbb{Z}^d \setminus \{0\}\} \le a \right\}. \tag{47}$$

2. *One particle might appear at different positions in the tree, i.e. the map $z : m \to \mathbb{T}^d \times \mathbb{R}^d$ might be not injective.*

The set $R(t,a)$, which can easily seen to be nonempty, is relevant due to periodic boundary conditions, which will lead to self-intersections of the cylinders. This happens, if $v - v_j$ is sufficiently close to a velocity v^*, where the components of $v^*_1, \dots, v^*_d$ are rationally dependent and $\eta v^* \in \mathbb{Z}^d$ with $\eta \in [-t,t] \setminus \{0\}$. The effect is not present in a setting where $(u,v) \in \mathbb{R}^d \times \mathbb{R}^d$.
The second effect is caused by the notorious recollisions. Both effects disappear for finite t as the diameter a tends to zero.

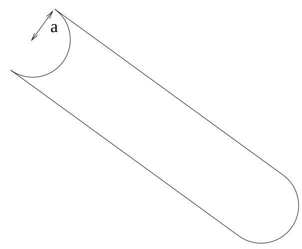

FIGURE 3. $C_l \cap \mathbb{T}^d \times \{v'\}$: The set of colliding initial data $C_l \subset \mathbb{T}^d \times \mathbb{R}^d$ consists for given $v' \in \mathbb{R}^d$ of a cylinder in $\mathbb{T}^d$ in direction $v' - v_l$ of length $s_l|v' - v_l|$ and diameter $2a$.

We stipulate now a strict order of the set of nodes m:

$$l < l' \text{ if either } |l| < |l'| \text{ or } (|l| = |l'| \text{ and } \bar{l} < \bar{l}') \text{ or } (\bar{l} = \bar{l}' \text{ and } l_{|l|} < l'_{|l|}). \tag{48}$$

This order is induced by the link between the collision time and the indices $l \in m$ in Definition 8.
Motivated by Remark 17 we define the set of "good" trees.

Definition 18. *For each $a_0 > 0$ the set of "good" trees $\mathcal{G}(a_0) \subset \mathcal{MT}$ consists of those trees $(m, \phi) \in \mathcal{MT}$ with the property that for all $0 < a \leq a_0$ and all $l \in m$*

$$v_l - v_{\bar{l}} \in \mathbb{R}^d \setminus R(t,a) \qquad \textit{(all parent-child-pairs are non-resonant),} \tag{49}$$

$$z_l \notin \cup_{\substack{l' < l \\ l' \neq \bar{l}}} C_{l'} \qquad \textit{(no particle appears twice in the tree),} \tag{50}$$

where we associate to each node $l \in m$ the set of colliding initial values

$$C_l = \left\{ z' \in \mathbb{T}^d \times \mathbb{R}^d \;\middle|\; \min_{s' \in [0, s_l]} \operatorname{dist}(z_l, z', s') \leq a \right\}, \tag{51}$$

and dist *as in (9) ignores overlap in the initial data.*

Note that $\mathcal{G}(a_0) \subset \mathcal{MT}$ is a family of sets which decreases with a_0. An elementary calculation yields that for all $v' \in \mathbb{R}^d \setminus (v_l + R(t,a))$

$$\mathcal{H}^d \left(C_l \cap (\mathbb{T}^d \times \{v'\}) \right) = \frac{\kappa_d}{n} |v_l - v'| s_l. \tag{52}$$

The concept of good trees will now be used to derive a more explicit characterization of the distributions $\hat{P}_k(\cdot \mid \Phi_{k-1})$.
As an intermediate step we recall a formula which yields the probability of certain complex events with respect to Poisson-point processes. Let $A \subset \cup_{N=0}^{\infty} (\mathbb{T}^d \times \mathbb{R}^d)^N$ be a symmetric set, i.e. $z \in A \cap (\mathbb{T}^d \times \mathbb{R}^d)^N$ if and only if $(z_{\pi(1)}, \dots, z_{\pi(N)}) \in A \cap (\mathbb{T}^d \times \mathbb{R}^d)^N$ for all permutations $\pi \in S_N$, where S_N is the symmetric group. We use the convention that $(\mathbb{T}^d \times \mathbb{R}^d)^0$ is a single point. For each realization $\omega \subset \mathbb{T}^d \times \mathbb{R}^d$ of the point process we chose an arbitrary enumeration of the elements of ω such that $\omega = \{z_1, \dots, z_N\}$. We say that $\omega \in A$ if $(z_1, \dots, z_N) \in A$; the choice of the enumeration is irrelevant since A is symmetric. It can be checked that if ω is a realization of the Poisson-point process with intensity $\mu \in M_+(\mathbb{T}^d \times \mathbb{R}^d)$, then

$$\operatorname{Prob}_{\mathrm{ppp}}(\omega \in A) = e^{-\mu(\mathbb{T}^d \times \mathbb{R}^d)} \sum_{N=0}^{\infty} \frac{1}{N!} \int_{A \cap (\mathbb{T}^d \times \mathbb{R}^d)^N} \mathrm{d}\mu(z_1) \dots \mathrm{d}\mu(z_n), \tag{53}$$

where the value of integral for $N = 0$ is 1 if $(\mathbb{T}^d \times \mathbb{R}^d)^0 \subset A$ and 0 else. By the definition of Poisson-point processes each set $\mathcal{C} \subset \mathbb{T}^d \times \mathbb{R}^d$ defines a projection denoted by $\mathcal{C} \cap \omega$. We recall the following fundamental independence-principle of Poisson-point processes which asserts that even if we have obtained a certain amount of information over a realization ω of a Poisson-point process it is still possible to use a suitably modified version of formula (53).

Lemma 19. *Let the random set $\omega \subset \mathbb{T}^d \times \mathbb{R}^d$ be distributed according to a Poisson point-process with density μ, $\bar{\mathcal{C}}, \mathcal{C} \subset \mathbb{T}^d \times \mathbb{R}^d$ and $A \subset \cup_{r=0}^{\infty} (\mathcal{C} \setminus \bar{\mathcal{C}})^r$ be symmetric. Then we obtain the following formula for the conditional probability of the event A:*

$$\text{(54)} \qquad \mathrm{Prob}_{\mathrm{ppp}} \left(\omega \cap \mathcal{C} \in A \;\middle|\; \omega \cap \bar{\mathcal{C}} = \emptyset \right) = \exp\left(-\mu(\mathcal{C} \setminus \bar{\mathcal{C}}) \right) \sum_{r=0}^{\infty} \frac{1}{r!} \int_{A \cap \mathcal{C}^r} \mathrm{d}\mu^r(z),$$

where $\mu^r = \underbrace{\mu \otimes \ldots \otimes \mu}_{r \text{ terms}}$.

Proof. See section 4. $\square$

To apply Lemma 19 we have to work with the phase space representation of trees. We use the decomposition $\Omega = \dot{\bigcup}_{m \in \mathcal{T}} \mathcal{E}(m) \cap \Omega$ and restrict our attention to $\Omega \subset \mathcal{E}(m)$ for some $m \in \mathcal{T}$. When we will apply eq. (54) to a given tree $m \in \mathcal{T}$, the number of points in $\omega \cap \mathcal{C}$ will be determined by m. Hence $A \subset (\mathcal{C} \setminus \bar{\mathcal{C}})^r$ for one r only which simplifies (54) to a single nontrivial term.

Note that for a general tree $\Phi = (m, \phi) \in \mathcal{MT}$ the number of nodes $\#m$ can be bigger than the number of particles involved in the collisions, i.e. it is possible that the map $z : m \to \mathbb{T}^d \times \mathbb{R}^d$ is not injective and $z_l = z_{l'}$ for some pair $l, l' \in m$, $l \neq l'$. This scenario corresponds to a bad tree where two nodes represent the same particle, see (50). For this reason we restrict our attention to sets Ω which are subsets of $\mathcal{G}(a)$. The excluded set has nonzero probability, however we will show that the probability of $\mathcal{MT} \setminus \mathcal{G}(a)$ tends with a to 0. By construction for all trees in $\Omega \subset \mathcal{G}(a)$ the map $l \mapsto z_l$ is injective.

The order defined by (48) for the nodes $l \in m$ induces a representation of the events $\Omega \subset \mathcal{E}(m)$ in phase-space coordinates, by

$$(z_l)_{l \in m} = (z_{l_1}, \ldots, z_{l_{\#m}})_{l_1, \ldots, l_{\#m} \in m} \text{ such that } l_1 < l_2 < \ldots < l_{\#m} \text{ in the order (48)}.$$

These events are denoted as

$$A(\Omega) = \{ (z_l)_{l \in m} | (m, \phi) \in \Omega \} \subset (\mathbb{T}^d \times \mathbb{R}^d)^{\#m}.$$

In the same spirit one obtains a one-to-one correspondence between the initial values of particles associated with the tree-nodes at height k and subsets of $(\mathbb{T}^d \times \mathbb{R}^d)^{\#m \cap \mathbb{N}^k}$:

$$Z_k = (z_l)_{|l|=k} \in (\mathbb{T}^d \times \mathbb{R}^d)^{\#(m \cap \mathbb{N}^k)}.$$

We will also need the conditional events

$$A_k(\Omega, \Phi) = \left\{ Z_k \in (\mathbb{T}^d \times \mathbb{R}^d)^{\#(m \cap \mathbb{N}^k)} \,|\, (Z_k, \Phi) \in \Omega \right\},$$

where $\Phi \in \mathcal{MT}_{k-1}$ and $(Z_k, \Phi) \in \mathcal{MT}_k$ is the tree obtained by attaching the leaves Z_k to the topmost nodes of Φ.

Recall that the density of the Poisson-point process which generates the initial positions of the particles is given by μ where

$$\int \mathrm{d}\mu(z)\, \varphi(z) = n \int_{\mathbb{R}^d} \mathrm{d}f_0(v) \int_{\mathbb{T}^d} \mathrm{d}u\, \varphi(u, v)$$

for every testfunction $\varphi \in C_c(\mathbb{T}^d \times \mathbb{R}^d)$.

Before applying Lemma 19 we have to specify the sets $\mathcal{C}$ and $\bar{\mathcal{C}}$. Fix $a_0 > 0$ and let $\Phi \in \mathcal{MT} \cap \mathcal{G}(a_0)$. We are interested in the distribution of those trees which coincide with

Φ up to level k. Clearly, the initial positions of the particles at height $k+1$ are contained in the set (compare Fig. 3)

$$\mathcal{C}_k(\Phi) := \bigcup_{l \in m \cap \mathbb{N}^k} C_l(\phi) \subset \mathbb{T}^d \times \mathbb{R}^d, \tag{55}$$

with $C_l(\phi) = C_l$ as in (51) and with $\Phi = (m, \phi)$. In order to apply formula (54) we have to identify the conditioning of the distribution $\omega \cap \mathcal{C}_k(\Phi)$. Define the collection of cylinders

$$\bar{\mathcal{C}}_k(\Phi) := \bigcup_{|l|<k} C_l(\phi) \subset \mathbb{T}^d \times \mathbb{R}^d \tag{56}$$

which contains those initial values that would affect the lower nodes. By construction the information on the point process ω that we have accumulated so far is given by $\omega \cap \bar{\mathcal{C}}_k(\Phi) = \{z_l \mid |l| \leq k\}$. Furthermore, since $\Phi \in \mathcal{G}(a_0)$ we have that $\omega \cap \mathcal{C}_k(\Phi) \cap \bar{\mathcal{C}}_k(\Phi) = \emptyset$. This implies that for each $\Omega \subset \mathcal{MT} \cap \mathcal{G}(a_0)$ and $\Phi \in \mathcal{MT}_k \cap \mathcal{G}(a_0)$

$$\hat{P}_{k+1}(\Omega \,|\, \Phi) = \mathrm{Prob}_{\mathrm{tppp}}(\mathcal{C}_k(\Phi) \cap \omega \in \mathrm{sym}(A_k(\Omega, \Phi)) \mid \mathcal{C}_k(\Phi) \cap \bar{\mathcal{C}}_k(\Phi) \cap \omega = \emptyset).$$

where $\mathrm{sym}(A)$ is the symmetrization of the set A, i.e. $(z_1, \ldots, z_N) \in \mathrm{sym}(A)$ if there exists a permutation $\pi \in S_N$ such that $(z_{\pi(1)}, \ldots, z_{\pi(N)}) \in A$; in particular $A \subset \mathrm{sym}(A)$. This is the crucial step where the complicated dependency on the past of the many-body evolution is reduced to a simple conditional expectation of the Poisson point process. Since $A(\Omega, \Phi) \cap \underbrace{\bar{\mathcal{C}}_k(\Phi) \times \ldots \times \bar{\mathcal{C}}_k(\Phi)}_{r \text{ terms}} = \emptyset$ for each r we can use formula (54) and deduce that

$$\hat{P}_{k+1}(\Omega \,|\, \Phi) = e^{-\hat{\Gamma}_k(\Phi)} \frac{1}{r!} \int_{\mathrm{sym}(A_{k+1}(\Omega,\Phi))} \mathrm{d}\mu^r(Z_{k+1}),$$

where

$$\hat{\Gamma}_k(\Phi) = \mu(\hat{\mathcal{C}}_k(\Phi)) \tag{57}$$

and

$$\hat{\mathcal{C}}_k(\Phi) = \mathcal{C}_k(\Phi) \setminus \bar{\mathcal{C}}_k(\Phi). \tag{58}$$

We use the convention that the value of the integral over $(\mathbb{T}^d \times \mathbb{R}^d)^0$ is 1.
As explained directly after Definition 8, each permutation of the labels $l \in m$ destroys the tree structure. Hence we obtain that if $z_\pi \in A$ and $z \in A$, then necessarily π is the identity transformation, i.e. $z_\pi = z$. This implies that if we replace in the above formula $\mathrm{sym}(A)$ by the non-symmetric uniquely ordered set A we have to drop the term $\frac{1}{r!}$.

$$\hat{P}_{k+1}(\Omega \,|\, \Phi) = e^{-\hat{\Gamma}_k(\Phi)} \int_{A_{k+1}(\Omega,\Phi)} \mathrm{d}\mu^r(Z_{k+1}). \tag{59}$$

Plugging the expression (59) for the conditional expectation $\hat{P}_{k+1}(\cdot \,|\, \Phi)$ into eq. (45) yields a representation of $\hat{P}_k$

Lemma 20. *Let $\Omega \subset \mathcal{G}(a) \cap \mathcal{T}_k$ be a Borel set, then*

$$\hat{P}_k(\Omega) = \sum_{m \in \mathcal{T}_k} \int_{A(\Omega)} \mathrm{d}\mu^{\#m}(z)\, e^{-\sum_{j<k} \hat{\Gamma}_j(\Phi(z))}. \tag{60}$$

The proof is given in section 4.
We return now to the collision representation of the trees. This means that the variables $(z_l)_{l\in m}$ in the integration are replaced by $(u_1, v_1) \times (s_l, \nu_l, v_l)_{l\in m\setminus\{1\}}$. The determinant of the derivative of this transformation is given by

$$\det D_\Phi z(\Phi) = \prod_{l\in m\setminus\{1\}} \left(a^{d-1}(\nu_l \cdot (v_l - v_{\bar{l}}))_+\right).$$

Thus changing coordinates in the integrals we obtain that for each $m \in \mathcal{T}$

$$\begin{aligned}
&\int_{A(\Omega)} e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi(z))}\, \mathrm{d}\mu^{\# m}(z) \\
&= \int_\Omega \mathrm{d}P_1(z_1)\, e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi)} \prod_{l\in m\setminus\{1\}} \left(n\, \mathrm{d}f_0(v_l)\, \mathrm{d}\nu_l\, \mathrm{d}s_l\, \chi_{[0,s_{\bar{l}}]}(s_l)\, a^{d-1}\left[(v_l - v_{\bar{l}})\cdot \nu_l\right]_+\right) \\
&\overset{(2)}{=} \int_\Omega \mathrm{d}P_1(z_1)\, e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi)} \prod_{l\in m\setminus\{1\}} \left(\mathrm{d}f_0(v_l)\, \mathrm{d}\nu_l\, \mathrm{d}s_l\, \chi_{[0,s_{\bar{l}}]}(s_l)\left[(v_l - v_{\bar{l}})\cdot \nu_l)\right]_+\right) \\
&= \int_\Omega \mathrm{d}\lambda^m(\phi)\, e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi)}.
\end{aligned}$$

Thus we have shown that for all $\Omega \subset \mathcal{G}(a)$

$$\hat{P}_k(\Omega) = \sum_{m\in\mathcal{T}_k} \int_{\Omega\cap\mathcal{E}(m)} e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi)}\, \mathrm{d}\lambda^m(\phi). \tag{61}$$

and

$$P_k(\Omega) = \hat{P}_k(\Omega) + e_k(\Omega), \tag{62}$$

where by eq. (24) the error has the form

$$e_k(\Omega) = \sum_{m\in\mathcal{T}_k} \int_{\Omega\cap\mathcal{E}(m)} \mathrm{d}\lambda^m(\phi) \left(e^{-\sum_{j<k}\Gamma_j(\Phi)} - e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi)}\right). \tag{63}$$

Lemma 21. *The error function $e_k(\Omega)$ in (63) is a non-negative measure.*

Proof. With (25), we first observe because $\Omega \subset \mathcal{G}(a)$

$$\Gamma_k(\Phi) = \sum_{l\in m, |l|=k} \gamma_l(\Phi) = \sum_{l\in m, |l|=k} \mu(C_l).$$

Then (55) and (58) imply

$$\Gamma_j(\Phi) \geq \mu(\mathcal{C}_k(\Phi)) \geq \mu(\hat{\mathcal{C}}_k(\Phi)) = \hat{\Gamma}_k(\Phi),$$

which implies the lemma. □

The last lemma shows that the finite size effects in the empirical distribution are due to intersections of the cylinders of colliding initial data. These effect can only decrease the collision rate. Formula (63) is the key for quantifying the difference between P_k and $\hat{P}_k$.

2.6. Total variation estimate of $P_k - \hat{P}_k$.

Proposition 22 (Tightness)**.** *Let $\mathcal{G}(a)$ the set of good trees from Definition 18, and $\Omega \subset \mathcal{G}(a_0)$. Then the following equations are true:*

$$\lim_{a_0\to 0} \inf_k P_k(\mathcal{G}(a_0)) = 1, \tag{64}$$

$$\lim_{a\to 0} \left(\sup\left\{ \left|\hat{P}_k(\Omega) - P_k(\Omega)\right| \;\middle|\; k > 0, \Omega \subset \mathcal{G}(a_0)\,\mathrm{Borel}\right\}\right) = 0 \text{ if } a_0 \text{ is fixed.} \tag{65}$$

The assertions of the proposition amount to establishing convergence of $\hat{P}_k$ to P_k in the total-variation sense uniformly in k.
The proof relies on several simple, but somehow technical estimates and can be found at the end of the subsection. We will first estimate the size of the set $R(t,a)$.

Lemma 23. *Under the assumption of theorem 2*

$$\lim_{a\to 0}\int_{R(t,a)}(1+|v|)\,\mathrm{d}f_0(v)=0. \tag{66}$$

The proof can found in section 4.
For technical reasons we decouple the dependency of $\mathcal{G}$ and $\hat{P}_k$ on the scaling parameter a. We will construct a family of sets of trees $\hat{\mathcal{G}}(a)\subset\mathcal{G}(a)$ with the following two properties

$$\begin{aligned}
&\lim_{a_0\to 0}\inf_k P_k\left(\hat{\mathcal{G}}(a_0)\right)=1,\\
&\lim_{a_0\to 0}\lim_{a\to 0}\sup_{\Omega\subset\mathcal{MT},k\geq 1}\left|\hat{P}_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\right)-P_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\right)\right|=0.
\end{aligned}$$

The limit $a\to 0$ is relevant in the second formula through the dependence of $\hat{P}_k$ on a.
The idea is that the trees in the sets $\hat{\mathcal{G}}(a_0)$ have additional good properties which are controlled by a_0. It is quite clear that for our choice of $\hat{\mathcal{G}}(a_0)$ (see (67)) eq. (69) holds even for fixed a_0 but without the limit the proof becomes more complicated.
Now we construct $\hat{\mathcal{G}}$. It is the intersection of good trees for various a. To compare these we only consider the collision representation $\mathbb{R}^d\times S^{d-1}\times\mathbb{R}^+$ of trees, which is independent of a, while the initial position of the colliding particles u_l varies with a.

Definition 24. *Let* $\varepsilon(a)$ *and* $V(a)$ *be monotone positive functions of* a *such that* $\lim_{a\to 0}\varepsilon(a)=0$ *and* $\lim_{a\to 0}V(a)=+\infty$. *We define the set*

$$\hat{\mathcal{G}}(a_0)=\bigcap_{0<a<a_0}\left\{(m,\phi)\in\mathcal{G}(a)\,\middle|\,\min_{l\in m}|v_l-v_{\bar{l}}|\geq\varepsilon(a)\text{and}|v|\leq V(a)\right.$$
$$\left.\text{and}\min_{l\in m}\min_{l'<l,l'\neq\bar{l}}\left(1-\left|\frac{v_l-v_{\bar{l}}}{|v_l-v_{\bar{l}}|}\cdot\frac{v_{l'}-v_{\bar{l}}}{|v_{l'}-v_{\bar{l}}|}\right|\right)\geq\varepsilon(a)\right\}. \tag{67}$$

Lemma 25. *For any monotone* $V(.)$ *and* $\varepsilon(.)$ *in the definition 24 of* $\hat{\mathcal{G}}(a_0)$, *we have*

$$\lim_{a_0\to 0}\inf_k P_k\left(\hat{\mathcal{G}}(a_0)\right)=1. \tag{68}$$

Proof. The functions $\varepsilon(.)$ and $V(.)$ are monotone in a with $\lim_{a\to 0}\varepsilon(a)=0$ and $\lim_{a\to 0}V(a)=\infty$. The set $\hat{\mathcal{G}}(a_0)$ is monotonously decreasing in a_0, as we are using the collision data only in the nodes for the intersection of all $a<a_0$ and as $\varepsilon\searrow 0$ and $V\nearrow\infty$. By the monotone convergence theorem for sets, we obtain

$$\lim_{a\to 0}P_t(\hat{\mathcal{G}}(a))=P_t(\cup_{a>0}\hat{\mathcal{G}}(a)),$$

then $(m,\phi)\in\cup_{a>0}\hat{\mathcal{G}}(a)$ if for all $l\in m$:

$$\begin{aligned}
v_l-v_{\bar{l}}&\notin R(t,0),\\
z_l&\notin\cup_{\substack{l'<l\\ l'\neq\bar{l}}}C_{l'}(a=0).
\end{aligned}$$

Using (6), we see that for any given skeleton m the probability of violating these conditions is zero. Thus we obtain $P_t(\mathcal{E}(m)\cap\cup_{a>0}\hat{\mathcal{G}}(a))=P_t(\mathcal{E}(m))$. Hence

$$\lim_{a\to 0}P_t(\hat{\mathcal{G}}(a))=\sum_{m\in\mathcal{T}}P_t(\mathcal{E}(m)\cap\cup_{a>0}\hat{\mathcal{G}}(a))=\sum_{m\in\mathcal{T}}P_t(\mathcal{E}(m))=1.$$

$\square$

Lemma 26. *Let $V(.)$ and $\varepsilon(.)$ be monotone functions in the definition 24 of $\hat{\mathcal{G}}(a_0)$, then*

$$\lim_{a_0\to 0}\lim_{a\to 0}\sup_{\Omega\subset\mathcal{MT}\text{Borel},k\geq 1}\left|\hat{P}_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\right)-P_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\right)\right|=0. \tag{69}$$

Proof. Fix a_0 and let $\Omega\subset\hat{\mathcal{G}}(a_0)$. We first split off the contribution of the trees with many nodes. By Lemma 13, the expected value of the number of nodes $\#m$ in a tree m is bounded by $K_{\text{ini}}\exp(\kappa_d K_{\text{ini}}t)+1$. As $\#m$ is a positive function, we can use Markov's inequality and the estimate on the expected value of nodes to obtain the estimate

$$\sum_{\#m-1>r}P_k(\mathcal{E}(m))=P_k(X(m)>r)<\frac{\mathbb{E}(X)}{r}\leq\tfrac{K_{\text{ini}}}{r}\exp(\kappa_d K_{\text{ini}}t). \tag{70}$$

This estimate gives us control over the error which arises if we ignore all trees with more than r nodes:

$$\begin{aligned}1=\sum_{m\in\mathcal{T}}P_k(\mathcal{E}(m))&=\sum_{\substack{m\in\mathcal{T}\\ \#m-1\leq r}}P_k(\mathcal{E}(m))+\sum_{\substack{m\in\mathcal{T}\\ \#m-1>r}}P_k(\mathcal{E}(m))\\ &\leq\sum_{\substack{m\in\mathcal{T}\\ \#m-1\leq r}}P_k(\mathcal{E}(m))+\tfrac{K_{\text{ini}}}{r}e^{\kappa_d K_{\text{ini}}t}.\end{aligned} \tag{71}$$

In particular, if $r\geq\frac{K_{\text{ini}}}{\delta}e^{\kappa_d K_{\text{ini}}t}+1$, then

$$\sum_{\substack{m\in\mathcal{T}\\ \#m\leq r}}P_k(\mathcal{E}(m))\geq 1-\delta. \tag{72}$$

Denoting

$$\begin{aligned}I_1&=\sup_k\sum_{\substack{m\in\mathcal{T}\\ \#m\leq r}}\left|\hat{P}_k(\Omega\cap\mathcal{E}(m))-P_k(\Omega\cap\mathcal{E}(m))\right|,\\ I_2&=\sup_k\sum_{\substack{m\in\mathcal{T}\\ \#m>r}}\hat{P}_k\left(\hat{\mathcal{G}}(a)\cap\mathcal{E}(m)\right),\\ I_3&=\sup_k\sum_{\substack{m\in\mathcal{T}\\ \#m>r}}P_k\left(\hat{\mathcal{G}}(a)\cap\mathcal{E}(m)\right),\end{aligned}$$

then for each $r>0$ one obtains that

$$\lim_{a\to 0}\sup_k\left|\hat{P}_k(\Omega)-P_k(\Omega)\right|\leq\lim_{a\to 0}(I_1+I_2+I_3).$$

We will show that $\lim_{a\to 0}I_1=0$ and $\limsup_{a\to 0}(I_2+I_3)=o(1)$ as δ tends to 0 (cf. eq. (72)).
First we consider I_1. Since there is only a finite number of tree skeletons with at most r nodes it suffices to show that

$$\lim_{a\to 0}\sup_k\left|\hat{P}_k(\Omega\cap\mathcal{E}(m))-P_k(\Omega\cap\mathcal{E}(m))\right|=0$$

for each $m\in\mathcal{T}$ such that $\#m\leq r$. We have seen earlier (eq. (62)) that $P_k(\Omega\cap\mathcal{E}(m))=\hat{P}_k(\Omega\cap\mathcal{E}(m))+e(\Omega\cap\mathcal{E}(m))$ where

$$0\leq e(\Omega\cap\mathcal{E}(m))=\int_{\Omega\cap\mathcal{E}(m)}\mathrm{d}\lambda^m(\phi)\left(e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi)}-e^{-\sum_{j<k}\Gamma_j(\Phi)}\right).$$

Since $\Gamma_j(\Phi) \geq \hat{\Gamma}_j(\Phi)$ (cf. Lemma 21) one obtains

$$e(\Omega \cap \mathcal{E}(m)) \leq \int_{\hat{\mathcal{G}}(a)\cap\mathcal{E}(m)} \mathrm{d}\lambda^m(\phi)\, e^{-\sum_{j<k}\Gamma_j(\Phi)} \left(e^{\sum_{j<k}(\Gamma_j(\Phi)-\hat{\Gamma}_j(\Phi))} - 1\right).$$

We will demonstrate that there is a number $K(a_0, a) > 0$ such that $\lim_{a\to 0} K(a_0, a) = 0$ and for all $\Phi \in \mathcal{E}(m) \cap \hat{\mathcal{G}}(a_0)$ and all $j \in \mathbb{N}$ the estimate

$$0 \leq \Gamma_j(\Phi) - \hat{\Gamma}_j(\Phi) \leq K(a_0, a) \tag{73}$$

holds. Since $\int_{\mathcal{E}(m)} \mathrm{d}\lambda^m(\phi)\, e^{-\sum_{j<k}\Gamma_j(\Phi)} \leq 1$, and $\sum_{j<k} 1 \leq \#m \leq r$ this yields the bound

$$0 \leq P(\Omega \cap \mathcal{E}(m)) - \hat{P}(\Omega \cap \mathcal{E}(m)) \leq rK(a_0, a)e^{rK(a_0,a)}. \tag{74}$$

Thus estimate (73) implies $\lim_{a\to 0} I_1 = 0$. To prove (73) we recall that by definition (57), see also (55), (56) and (58)

$$\begin{aligned}
\hat{\Gamma}_j(\Phi) =& n \int_{\mathbb{R}^d} \mathrm{d}f_0(v')\, \mathcal{H}^d \left(\hat{\mathcal{C}}_j(\Phi) \cap (\mathbb{T}^d \times \{v'\})\right) \\
\geq& n \sum_{|l|=j} \int_{\mathbb{R}^d} \mathrm{d}f_0(v')\, \mathcal{H}^d \left(C_l(\phi) \cap (\mathbb{T}^d \times \{v'\})\right) - e_1 \\
=& n \sum_{|l|=j} \int_{\mathbb{R}^d \setminus (v_l + R(t,a))} \mathrm{d}f_0(v')\, \mathcal{H}^d \left(C_l(\phi) \cap (\mathbb{T}^d \times \{v'\})\right) - e_1 + e_2 \\
=& \sum_{|l|=j} \int_{\mathbb{R}^d \setminus (v_l + R(t,a))} \mathrm{d}f_0(v)\, \kappa_d |v_l - v| t - e_1 + e_2 = \Gamma_j(\Phi) - e_1 + e_2 + e_3,
\end{aligned}$$

where the error terms are defined as follows

$$\begin{aligned}
e_1 &= n \int_{\mathbb{R}^d} \mathrm{d}f_0(v')\, \mathcal{H}^d \left((\mathcal{C}_j(\Phi) \cap \bar{\mathcal{C}}_j(\Phi)) \cap (\mathbb{T}^d \times \{v'\})\right), \\
e_2 &= n \sum_{|l|=j} \int_{v_l + R(t,a)} \mathrm{d}f_0(v')\, \mathcal{H}^d \left(C_l(\phi) \cap (\mathbb{T}^d \times \{v'\})\right), \\
e_3 &= \sum_{|l|=j} \int_{R(t,a)} \mathrm{d}f_0(v)\, \kappa_d |v| t.
\end{aligned}$$

We set $K(a_0, a) = e_1 - e_2 - e_3$ and show that $\lim_{a\to 0} e_j = 0$ for $j = 1, 2, 3$. For all $v' \in \mathbb{R}^d$ one obtains that $n\mathcal{H}^d \left(C_l(\phi) \cap (\mathbb{T}^d \times \{v'\})\right) \leq \kappa_d |v_l - v'| t$ irrespective whether $v' \in v_l + R(t, a)$ or not, since it can be bounded by the length of the path, and intersections will only reduce the measure.
Hence, using that $\sum_{|l|=j} 1 \leq \#m \leq r$

$$e_2 + e_3 \leq 2\, \kappa_d\, r\, t \int_{R(t,a)} |v|\, \mathrm{d}f_0(v)$$

and eq. (66) yields that $\lim_{a\to 0}(e_2 + e_3) = 0$.
It remains to estimate e_1. This is the only part where estimates are not uniform and depend on the constants $\varepsilon(a_0) < 1$ and $V(a_0)$.
To bound $\mathcal{H}^d \left(C_{\bar{l}}(\phi) \cap C_{l'}(\phi) \cap (\mathbb{T}^d \times \{v'\})\right)$, we define for $|v'| \leq \bar{V}$ the number $c(a_0, a, v')$ to be the maximum volume contained within the intersection of two cylinders of diameter

a and axes $v-v'$ and $v-v''$ if v, v' and v'' are constrained in a certain geometrical way:

$$c(a_0, a, v') = \sup\Big\{\zeta(u', u'', v, v', v'', a) \,\Big|\, u', u'' \in \mathbb{T}^d, v, v'' \in \mathbb{R}^d,\ |v'-v''| \geq \varepsilon(a_0) \\ \text{and } |v|, |v''| \leq V(a_0) \text{ and } \left|\tfrac{v-v'}{|v-v'|} \cdot \tfrac{v''-v'}{|v''-v'|}\right| \leq 1-\varepsilon(a_0)\Big\},$$

where

$$\zeta(u, u', v, v', v'', a) = \mathcal{H}^d\Big(\Big\{u \in \mathbb{T}^d \,\Big|\, \inf_{s\in[0,t]} |u-u'+s(v-v')|_{\mathbb{T}^d} \leq a \\ \text{and } \inf_{s\in[0,t]} |u-u''+s(v-v'')|_{\mathbb{T}^d}\} \leq a\Big\}\Big).$$

The cylinders can intersect at most $((\bar{V}+V(a_0))t+1)^2$ times. The volume of each intersection is bounded from above by $(2a)^{d-1}\ell$ where ℓ is the maximal length of a line segment which is parallel to $v-v'$ and is contained in the cylinder with axis parallel to $v''-v'$. A simple geometric consideration yields that $\ell = \frac{2a}{|\sin\psi|}$, where ψ is the angle enclosed by the vectors $v-v'$ and $v''-v'$. The law of sines implies that $\sin(\psi) = \frac{|v'-v''|}{|v-v'|}\sin(\psi_0)$, where ψ_0 is the angle enclosed by $v-v''$ and $v'-v''$. Since $\cos(\psi_0) \leq 1-\varepsilon$ and $|v'-v''| \geq \varepsilon$ we obtain that $|\sin(\psi)| \geq \frac{1}{|v-v'|}\varepsilon^{\frac{3}{2}}$ and thus

$$\mathcal{H}^d\left(C_{\bar{l}}(\phi) \cap C_{l'}(\phi) \cap (\mathbb{T}^d \times \{v\})\right) \leq 2^d a^{d-1}\, a\,\varepsilon(a_0)^{-\frac{3}{2}}(\bar{V}+V(a_0))((\bar{V}+V(a_0))t+1)^2.$$

Using (2) and that there are less than $r^2/2$ possible pairs $(\bar{l}, l')$ we find that

$$\begin{aligned} e_1 =& n\int_{\mathbb{R}^d\setminus B_{\bar{V}}(0)} \mathrm{d}f_0(v')\,\mathcal{H}^d\left((\mathcal{C}_j(\Phi)\cap\bar{\mathcal{C}}_j(\Phi))\cap(\mathbb{T}^d\times\{v'\})\right) \\ &+ n\int_{B_{\bar{V}}(0)} \mathrm{d}f_0(v')\,\mathcal{H}^d\left((\mathcal{C}_j(\Phi)\cap\bar{\mathcal{C}}_j(\Phi))\cap(\mathbb{T}^d\times\{v'\})\right) \\ \leq& 2^d\, r^2\, a\,\varepsilon(a_0)^{-\frac{3}{2}}(\bar{V}+V(a_0))((\bar{V}+V(a_0))t+1)^2 + \int_{B_{\bar{V}}(0)} \mathrm{d}f_0(v')\kappa_d(|v'|+V(a_0)), \end{aligned}$$

by choosing $\bar{V}$ large, the last term is arbitrarily small uniformly in a. In particular $\lim_{a\to 0} e_1 = 0$ if a_0 is kept fixed. Thus we have shown that $\lim_{a\to 0} K(a_0, a) = 0$ and thereby $\lim_{a\to 0} I_1 = 0$, i.e. we have shown the convergence in (69) for finite trees of size less than r for any fixed a_0:

(75)

$$\lim_{a\to 0} \sup_{\substack{\Omega\subset\mathcal{MT}\text{ Borel}\\ k\geq 1}} \left|\hat{P}_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\cap\bigcup_{m\in\mathcal{T},\#m\leq r}\mathcal{E}(m)\right) - P_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\cap\bigcup_{m\in\mathcal{T},\#m\leq r}\mathcal{E}(m)\right)\right| = 0.$$

We finish the proof by showing that $\lim_{\delta\to 0}\lim_{a_0\to 0}\lim_{a\to 0}(I_2+I_3) = 0$. Equation (70) yields

$$I_3 = \sup_k \sum_{\substack{m\in\mathcal{T}\\ \#m>r}} P_k(\hat{\mathcal{G}}(a)\cap\mathcal{E}(m)) \leq \frac{1}{r-1}K_{\text{ini}}\exp(\kappa_d K_{\text{ini}}t) \leq \delta \tag{76}$$

and in a similar way we obtain

$$
\begin{aligned}
\lim_{a_0\to 0}\lim_{a\to 0} I_2 &= \lim_{a_0\to 0}\lim_{a\to 0}\sup_k \sum_{\substack{m\in\mathcal{T}\\ \#m>r}} \hat{P}_k\left(\hat{\mathcal{G}}(a)\cap\mathcal{E}(m)\right)\\
&\le \lim_{a_0\to 0}\lim_{a\to 0}\sup_k \hat{P}_k(\hat{\mathcal{G}}(a_0)) - \lim_{a_0\to 0}\lim_{a\to 0}\inf_k \sum_{\substack{m\in\mathcal{T}\\ \#m\le r}} \hat{P}_k\left(\hat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)\right)\\
&\overset{(68,75)}{=} 1 - \lim_{a_0\to 0}\inf_k \sum_{\substack{m\in\mathcal{T}\\ \#m\le r}} P_k\left(\hat{\mathcal{G}}(a_0)\cap\mathcal{E}(m)\right) \overset{(76)}{\le} \delta.
\end{aligned}
$$

Equation (68) yields that the last expression converges to 0 uniformly in a_0 as δ tends to 0. Thus we have demonstrated that (69) is satisfied. □

Now we are in the position to give the proof of Proposition 22.

Proof of Proposition 22. We show that (68) and (69) imply (65): Since $\hat{P}_k$ and P_k are probability measures, eq. (69) implies that

$$
\lim_{a_0\to 0}\lim_{a\to 0}\sup_k \left|\hat{P}_k\left(\mathcal{MT}\setminus\hat{\mathcal{G}}(a_0)\right) - P_k\left(\mathcal{MT}\setminus\hat{\mathcal{G}}(a_0)\right)\right| = 0. \tag{77}
$$

Let now $\Omega\subset\hat{\mathcal{G}}(a_0)$ for some $a_0>0$ and fix $\varepsilon>0$. Then

$$
\begin{aligned}
\lim_{a_0\to 0}\lim_{a\to 0}\sup_k |\hat{P}_k(\Omega)-P_k(\Omega)| &\le \lim_{a_0\to 0}\lim_{a\to 0}\sup_k \left|\hat{P}_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\right) - P_k\left(\Omega\cap\hat{\mathcal{G}}(a_0)\right)\right|\\
&\quad + \lim_{a_0\to 0}\lim_{a\to 0}\left(\sup_k \hat{P}_k\left(\Omega\setminus\hat{\mathcal{G}}(a_0)\right) + \sup_k P_k\left(\Omega\setminus\hat{\mathcal{G}}(a_0)\right)\right)\\
&\overset{(69)}{=} \lim_{a_0\to 0}\lim_{a\to 0}\sup_k \hat{P}_k\left(\Omega\setminus\hat{\mathcal{G}}(a_0)\right) + \lim_{a_0\to 0}\sup_k P_k\left(\Omega\setminus\hat{\mathcal{G}}(a_0)\right)\\
&\overset{(77)}{\le} 2\lim_{a_0\to 0}\sup_k P_k\left(\mathcal{MT}\setminus\hat{\mathcal{G}}(a_0)\right) \overset{(68)}{=} 0.
\end{aligned}
$$

Equation (64) follows directly from (68) since $\hat{\mathcal{G}}(a)\subset\mathcal{G}(a)$. □

2.7. **Proof of the main results.**

Proof of Theorem 2. We demonstrate that the distribution of a single tagged particle satisfies the Boltzmann equation. Let $A\subset\mathbb{T}^d\times\mathbb{R}^d$ be a Borel set and define $\Omega(A)\subset\mathcal{MT}$ by

$$
\Omega(A) = \{\Phi\in\mathcal{MT}\mid \beta_1(m)=1 \text{ and } z_1\in A\},
$$

which is a Borel set in $\mathcal{MT}$ (23), as $\beta_1(m)=1$ is a property of m alone. With this notation we obtain that for every $a_0>0$

$$
\begin{aligned}
&\left|\lim_{a\to 0}\lim_{k\to\infty}\hat{P}_{t,k}(\Omega) - \int_A \mathrm{d}u\,\mathrm{d}f_t(v)\right| \overset{\text{Lemma 5}}{=} \lim_{a\to 0}\lim_{k\to\infty}\left|\hat{P}_{t,k}(\Omega) - \int_A \mathrm{d}u\,\mathrm{d}f_{t,k-1}(v)\right|\\
&\overset{\text{Proposition 14}}{=} \lim_{a\to 0}\lim_{k\to\infty}\left|\hat{P}_{t,k}(\Omega) - P_{t,k}(\Omega)\right|\\
&\quad = \lim_{a\to 0}\lim_{k\to\infty}\left|\hat{P}_{t,k}(\Omega\cap\mathcal{G}(a_0)) - P_{t,k}(\Omega\cap\mathcal{G}(a_0)) - P_{t,k}(\Omega\setminus\mathcal{G}(a_0)) + \hat{P}_{t,k}(\Omega\setminus\mathcal{G}(a_0))\right|\\
&\quad \overset{(65)}{\le} \lim_{a\to 0}\lim_{k\to\infty} P_{t,k}(\mathcal{MT}\setminus\mathcal{G}(a_0)) + \lim_{a\to 0}\lim_{k\to\infty}\hat{P}_{t,k}(\mathcal{MT}\setminus\mathcal{G}(a_0)).
\end{aligned}
$$

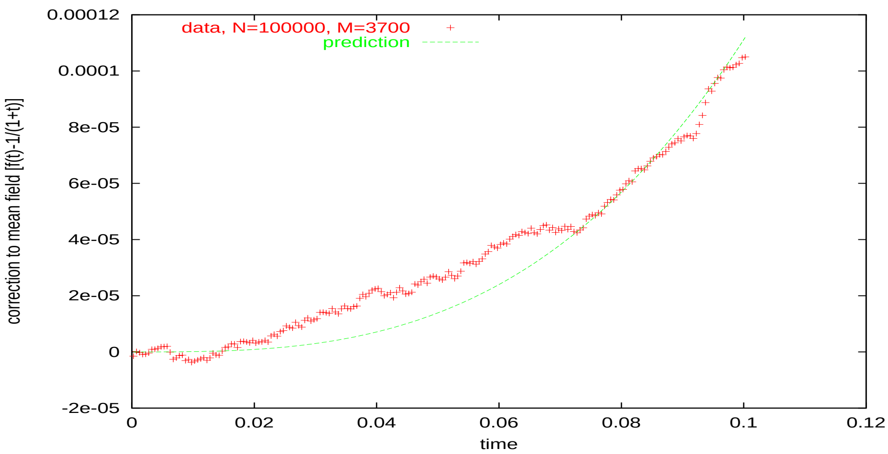

FIGURE 4. Comparison between the empirical probability of colliding and the idealized prediction. The dashed line is the cubic parabola $t \mapsto \frac{1}{9}t^3$, the signs '+' mark the difference between the number of non-collided particles at time t divided by n and the idealized prediction $\frac{1}{1+t}$.

Now using that $\hat{P}_{t,k}$ and $P_{t,k}$ are probability measures and eq. (65) again for $\tilde{\Omega} := \mathcal{MT} \cap \mathcal{G}(a_0)$, we obtain, for all $k \in \mathbb{N}$, that $\lim_{a\to 0} \hat{P}_{t,k}(\mathcal{MT} \setminus \mathcal{G}(a_0)) = 1 - \lim_{a\to 0} \hat{P}_{t,k}(\mathcal{G}(a_0)) = 1 - P_{t,k}(\mathcal{G}(a_0)) = P_{t,k}(\mathcal{MT} \setminus \mathcal{G}(a_0))$. Now proceeding

$$\leq 2 \lim_{k\to\infty} P_{t,k}(\mathcal{MT} \setminus \mathcal{G}(a_0)),$$

we send now a_0 to 0, apply (64) and obtain that $\lim_{a_0\to 0} \lim_{k\to\infty} P_{t,k}(\mathcal{MT} \setminus \mathcal{G}(a_0)) = 0$, hence $\lim_{a\to 0} \lim_{k\to\infty} \hat{P}_{t,k}(\Omega) = \int_A \mathrm{d}u \, \mathrm{d}f_t(v)$, and the proof of Theorem 2 is complete. □

Proof of Corollary 3. Equation (15), Lemma 5 and Remark 16 show that $\mathbf{1} \otimes f_{t,k}$, $\mathbf{1} \otimes f_t$ and $\hat{f}^{(a)}_{t,k}$ are absolutely continuous with respect to $\mathbf{1} \otimes f_0$. The calculation above implies convergence of $\hat{f}^{(a)}_{t,k}$ to $\mathbf{1} \otimes f_t$ in $L^1(\mathbf{1} \otimes f_0)$. □

3. THE EFFECT OF CONCENTRATIONS

We illustrate now that the idealized theory does not capture the many-particle dynamics if the initial distribution f_0 exhibits strong concentrations. To simplify the long calculations at the end of the proof we assume that $d = 2$, but similar results are expected to hold in the case $d = 3$.

Theorem 27. *Let $v \in \mathbb{R}^2$ be such that $|v| = \frac{1}{2}$ and set $f_0 = \frac{1}{2}(\delta(\cdot - v) + \delta(\cdot + v))$. If $\hat{Q}(t) = \lim_{a\to 0} \lim_{k\to\infty} \hat{P}_{t,k}(\beta_1 = 1)$ denotes the empirical probability that a tagged particle does not collide, then*

$$\lim_{t\to 0} \frac{1}{t^3} \left(\hat{Q}(t) - \int_{\mathbb{R}^2} \mathrm{d}f_t(v) \right) = \frac{1}{9}, \tag{78}$$

where $f_t = \frac{1}{1+t} f_0$ is the unique solution of the Boltzmann equation (11) which satisfies the initial condition $f_{t=0} = f_0$.

A numerical simulation (Fig. 3) illustrates the prediction (78).

Proof. We could use the tree measures P_t and $\hat{P}_t$ to prove the assertion. To keep the notation as simple as possible and focus on the essential computation we chose a slightly different approach based on Taylor expansion.

It can be assumed without loss of generality that the initial value of the tagged particle is $(0, v)$. We define the set

$$M_\lambda := \left\{ u \in \mathbb{T}^d \;\middle|\; \min_{s\in[0,t]} |2sv - u| \leq \rho \right\},$$

which is basically a cylinder with radius ρ and centerline given by the particle-trajectory without collisions and contains the initial positions of those particles that might collide with the tagged particle before time t. The parameter $\rho > 0$ is a function of λ such that $\mathrm{vol}(M_\lambda) = at\lambda$, i.e. ρ solves

$$\pi\rho^2 + 2\rho t - a\lambda t = 0. \tag{79}$$

The idea is that for short times the survival probability should be dominated by events where the number of initial positions which fall into the set M_λ is small. It turns out that the survival probability conditional to having j initial positions in M_λ can be computed explicitly provided that $\lambda \geq j + 1$. The reason is that for sufficiently large diameter the survival probability is independent of the configuration outside the cylinder.
Since only half of the particles are potential collision partners and the modulus of the relative velocity is 1 the collision rate is 1 (recall that $\kappa_2 = 2$). By construction of λ the probability that the total number of particles whose initial position is contained in M_λ equals k is given by $e^{-\lambda t}\frac{(\lambda t)^k}{k!}$.
Let $p_k(\lambda, t)$ be the probability that the particle does not collide before time t if there are precisely k particles contained in M_λ. We will show later that in the limit where a tends to 0 the probabilities p_k become independent of t. This is expected since all particles in M_λ, except those near the ends, will either collide or leave the cylinder before time t. For small a the cylinder is very slender and only little volume is contained in the caps near the ends. For this reason we will not show the dependency on t in future.

Lemma 28. *For all $j \in \mathbb{N}$*

$$\lim_{t\to 0} \frac{1}{t^j} \left| \hat{Q}(t) - e^{-\lambda t} \sum_{k=0}^{j} \tfrac{(\lambda t)^k}{k!} p_k(\lambda) \right| = 0. \tag{80}$$

Proof. Let $\omega = \{u_0(i), \mid i = 1 \dots N\}$ be the set of initial positions and $P_j = \mathrm{Prob}(\#(\omega \cap M_\lambda) > j)$ be the probability that M_λ contains more than j particles. Clearly

$$P_j = e^{-\lambda t}\left(e^{\lambda t} - \sum_{k=0}^{j} \frac{(\lambda t)^k}{k!} \right) \leq e^{-\lambda t} t^{j+1} \sup_{s\in[0,t]} \frac{\lambda^{j+1}}{(j+1)!} e^{\lambda s} = \frac{\lambda^{j+1}}{(j+1)!} t^{j+1},$$

where the inequality is due to Taylor's theorem. □

We will only be interested in the case $j = 3$.

Idealized behavior. Let $Q(t) := \frac{1}{1+t}$ be the particle density predicted by the idealized theory. We are seeking idealized probabilities $p_k^{\mathrm{id}}(\lambda) \in [0, 1]$ such that

$$Q(t) = e^{-\lambda t} \sum_{k=0}^{\infty} \tfrac{(\lambda t)^k}{k!} p_k^{\mathrm{id}}(\lambda). \tag{81}$$

Replacing the exponential function in (81) by the power series one obtains that

$$\sum_{l,m=0}^{\infty} \frac{(-\lambda t)^l}{l!} \frac{(\lambda t)^m}{m!} p_m^{\mathrm{id}}(\lambda) = \sum_{k=0}^{\infty} (-t)^k. \tag{82}$$

Ordering the left hand side by powers of t and equating coefficients yields the following hierarchical set of equations for the probabilities p_k

$$\sum_{l=0}^{k} \frac{(-1)^l}{l!(k-l)!} p^{\text{id}}_{k-l} = \left(-\frac{1}{\lambda}\right)^k .$$

We can use the equations above to determine p^{id}_k recursively and obtain that

$$p^{\text{id}}_k = (-1)^k \frac{k!}{\lambda^k} - \sum_{l=1}^{k} (-1)^l \binom{k}{l} p^{\text{id}}_{k-l}.$$

The recurrence relation can be solved explicitly and we obtain

$$p^{\text{id}}_k = \sum_{l=0}^{k} \frac{k!}{(k-l)!} \left(-\frac{1}{\lambda}\right)^l . \tag{83}$$

Equation (80) and (81) implies that if p_k does not agree with formula (83), then $\hat{Q}(t) \neq Q(t)$ if t is sufficiently small.

Computation of the empirical probabilities p_k. If $\lambda \geq k+1$ the probability p_k can be computed explicitly. The reason is that the diameter of the cylinder is so large that the collision probability is not influenced by the initial configuration outside M_λ and so small that the probability of initial configurations with overlap is negligible. To keep the notation as simple as possible we will from now on ignore errors coming from the finiteness of a and assume that the particles are intervals with length a perpendicular to the vector v. Explicit estimates of the dependency of p_2 on a are provided at the end of the proof, no approximation is involved in the case of p_1. The dependency of p_3 on a can be estimated analogously.
We will show now that for all $\lambda \geq 4$ the values of $p_k(\lambda)$, $k = 0, 1, 2, 3$ are given by $p_0 = 1$, $p_1 = 1 - \frac{1}{\lambda}$, $p_2 = 1 - \frac{2}{\lambda} + \frac{2}{\lambda^2}$, $p_3 = 1 - \frac{3}{\lambda} + \frac{6}{\lambda^2} - \frac{6}{\lambda^3} + \frac{\alpha_2}{\lambda^3}$ with $\alpha_2 = \frac{2}{3}$. This implies that

$$\lim_{t \to 0^+} \frac{\hat{Q}(t) - \frac{1}{1+t}}{t^3} = \frac{\alpha_2}{6} = \frac{1}{9}$$

and thus the claim.
Let $k \in \{0, 1, 2, 3\}$ be the number of particles contained in the set M_λ. For the sake of simplicity we say that the particles with velocity v are white and the particles with velocity $-v$ are black. One obtains 2^k different color distributions, each of those cases has the same probability of occurring.
We are now in a position to compute an explicit formula for the values of $p_k(\lambda)$. We have to consider several cases, depending on the direction and relative position of the particles in the path of the tagged particle. Particles traveling in the same direction as the tagged particle are denoted by w, particle in the other direction by b. The ordering of the particles in the cylinder is given in the index.

Computation of p_0.
It is clear that $p_0 = 1$ since there is no obstacle in M_λ.

Computation of p_1.
$p^w_1 = 1$,
$p^b_1 = 1 - \frac{2}{\lambda}$.
We obtain the overall probability $p_1 = \frac{1}{2}(p^w_1 + p^b_1) = 1 - \frac{1}{\lambda}$.

Computation of p_2.
$p_2^{ww} = 1$ (No collision possible),
$p_2^{bb} = (1 - \frac{2}{\lambda})^2$ (Probability of avoiding two independent black particles),
$p_2^{bw} = 1 - \frac{2}{\lambda}$ (Probability of avoiding one black particle, the position of the white particle is irrelevant),
$p_2^{wb} = 1 - \frac{2}{\lambda}(1 - \frac{2}{\lambda})$ (Probability of avoiding a black particle which might be removed by a white particle before it comes to a collision).
Adding the probabilities yields that $p_2 = \frac{1}{4}(p_2^{ww} + p_2^{bb} + p_2^{wb} + p_2^{bw}) = 1 - \frac{2}{\lambda} + \frac{2}{\lambda^2}$.

Computation of p_3.
$p_3^{www} = 1$ (No collision possible),
$p_3^{bbb} = (1 - \frac{2}{\lambda})^3$ (Probability of avoiding 3 independent black particles),
$p_3^{bww} = 1 - \frac{2}{\lambda}$ (Probability of avoiding 1 black particle, the white particles are irrelevant),
$p_3^{wbw} = 1 - \frac{2}{\lambda}(1 - \frac{2}{\lambda})$ (Probability of avoiding one black particle which might be removed by one white particle. The second white particle is irrelevant).
$p_3^{wwb} = 1 - \frac{2}{\lambda}(1 - \frac{2}{\lambda})^2$ (Probability of avoiding one black particle which might be removed by two independent white particles).
$p_3^{bbw} = (1 - \frac{2}{\lambda})^2$ (Probability of avoiding 2 independent black particles, the white particle is irrelevant).
$p_3^{bwb} = (1 - \frac{2}{\lambda})(1 - \frac{2}{\lambda}(1 - \frac{2}{\lambda}))$ (Probability of avoiding 2 independent black particles, the second black particle might be removed by a white particle).
$p_3^{wbb} = 1 - \frac{4}{\lambda} + \frac{12}{\lambda^2} - \frac{24}{\lambda^3} + 8\frac{\alpha_2}{\lambda^3}$
To demonstrate that the formula above indeed yields the correct value of p_3^{wbb} we introduce the coordinates perpendicular to v of the three particles $u_i \in \mathbb{R}$, $i = 1, 2, 3$ and consider four mutually exclusive scenarios. In three scenarios the probability of being scattered can be computed analogously to the preceding cases. As these computations are independent of a, we let $a = 1$ for notational convenience. For the particle to exist, we need that black particles in the cylinder (with either $|u_2| \leq 1$ or $|u_3| \leq 1$) have to be removed by the first white one.

$\mathrm{Prob}(|u_2| > 1 \text{ and } |u_3| > 1) = (1 - \frac{2}{\lambda})^2$, (Probability of avoiding 2 independent black particles, the white particle u_1 is irrelevant)
$\mathrm{Prob}(|u_2| \leq 1 \text{ and } |u_2 - u_1| \leq 1 \text{ and } |u_3| > 1) = \frac{4}{\lambda^2}(1 - \frac{2}{\lambda})$, (The white particle removes the first black particle, the second black particle is avoided.)
(The case $|u_2| \leq 1, |u_1 - u_2| > 1, |u_3| > 1$ does not contribute to p_3^{wbb} as it implies a collision.)
The last case is that the first black particle is avoided, while the white particle removes the second black particle, i.e. this case is $|u_2| > 1$, $|u_3| \leq 1$, $|u_1 - u_3| \leq 1$ and $|u_1 - u_2| > 1$. We split this into subcases: The first case is
$\mathrm{Prob}(|u_2| \geq 3 \text{ and } |u_3| \leq 1 \text{ and } |u_1 - u_3| \leq 1) = (1 - \frac{6}{\lambda})\frac{4}{\lambda^2}$.
To compute the probability of being scattered in the remaining case where $|u_2| \in (1, 3]$, $|u_3| \leq 1$, $|u_1 - u_3| \leq 1$ and $|u_1 - u_2| > 1$ we have to do an explicit integration.

$$\begin{aligned}
I_2 = &\int_{-1}^{1} \mathrm{d}u_3 \int_{u_3-1}^{u_3+1} \mathrm{d}u_1 \int_{1}^{3} \mathrm{d}u_2 \, (1 - \chi_{[-1,+1]}(u_1 - u_2)) \\
&+ \int_{-1}^{1} \mathrm{d}u_3 \int_{u_3-1}^{u_3+1} \mathrm{d}u_1 \int_{-3}^{-1} \mathrm{d}u_2 \, (1 - \chi_{[-1,+1]}(u_1 - u_2)) \\
=&2 \int_{-1}^{1} \mathrm{d}u_3 \int_{u_3-1}^{u_3+1} \mathrm{d}u_1 \int_{1}^{3} \mathrm{d}u_2 \, (1 - \chi_{[-1,+1]}(u_1 - u_2)),
\end{aligned}$$

as the other integral is obtained by the transformation $(u_1, u_2, u_3) \mapsto (-u_1, -u_2, -u_3)$. A simple but lengthy calculation yields that $I_2 = \frac{40}{3}$. The details of this calculation are irrelevant, but for the purpose of checking that this number is indeed correct the detailed calculations are included below. We obtain that

$$p_3^{wbb} = (1 - \tfrac{2}{\lambda})^2 + \tfrac{4}{\lambda^2}(1 - \tfrac{2}{\lambda}) + (1 - \tfrac{6}{\lambda})\tfrac{4}{\lambda^2} + \tfrac{I_2}{\lambda^3}.$$

Altogether this yields

$$\begin{aligned} p_3 =& \tfrac{1}{8}(p_3^{www} + p_3^{wwb} + p_3^{wbw} + p_3^{bww} + p_3^{wbb} + p_3^{bwb} + p_3^{bbw} + p_3^{bbb}) \\ =& 1 - \tfrac{3}{\lambda} + \tfrac{6}{\lambda^2} - \tfrac{6}{\lambda^3} + \tfrac{I_2 - 8}{8\lambda^3}, \end{aligned}$$

and therefore $\alpha_2 = \frac{I_2 - 8}{8} = \frac{2}{3}$.
We calculate now the value of I_2.

$$\begin{aligned} I_2 =& 2\int_{-1}^{1} \mathrm{d}u_3 \int_{u_3-1}^{0} \mathrm{d}u_1 \int_{1}^{3} \mathrm{d}u_2 \underbrace{(1 - \chi_{[-1,1]}(u_1 - u_2))}_{=1} \\ &+ 2\int_{-1}^{1} \mathrm{d}u_3 \int_{0}^{u_3+1} \mathrm{d}u_1 \int_{1}^{3} \mathrm{d}u_2 \, (1 - \chi_{[-1,1]}(u_1 - u_2)) \\ =& 4\int_{-1}^{1} \mathrm{d}u_3 \, (1 - u_3) + 2\int_{-1}^{1} \mathrm{d}u_3 \int_{0}^{u_3+1} \mathrm{d}u_1 \int_{1}^{3} \mathrm{d}u_2 \\ &- 2\int_{-1}^{1} \mathrm{d}u_3 \int_{0}^{u_3+1} \mathrm{d}u_1 \int_{1}^{3} \mathrm{d}u_2 \, \chi_{[-1,1]}(u_1 - u_2) \\ =& 8 + 8 - 2\int_{-1}^{1} \mathrm{d}u_3 \int_{0}^{u_3+1} \mathrm{d}u_1 \int_{\max(1, u_1-1)}^{1+u_1} \mathrm{d}u_2 \\ =& 16 - 2\int_{-1}^{1} \mathrm{d}u_3 \int_{0}^{u_3+1} \mathrm{d}u_1 \int_{1}^{1+u_1} \mathrm{d}u_2 \\ =& 16 - 2\int_{-1}^{1} \mathrm{d}u_3 \int_{0}^{u_3+1} \mathrm{d}u_1 \, u_1 = 16 - 2\int_{-1}^{1} \mathrm{d}u_3 \frac{1}{2}(u_3 + 1)^2 \\ =& 16 - \frac{1}{3}[(u_3 + 1)^3]_{u_3=-1}^{u_3=1} = 16 - \frac{8}{3}. \end{aligned}$$

We provide now an explicit estimate of the dependency of p_2 on a. Due to our choice of ρ (79) and as $\lambda \geq 4$ there are not any finite size effects for p_2^{ww}, p_2^{bb} and p_2^{bw}. Only p_2^{wb} depends on a: The probability that the white particle removes the black one, that will be hit at a time $0 \leq s \leq t$, is given by the probability of finding an extra white particle in M_λ such that it hits before time s. This probability is given by the volume quotient: $\frac{2as}{\pi\rho^2/2 + 2s\rho}$. Integrating this yields

$$p_2^{wb} = 1 - \left(1 - \frac{2}{\lambda}\right) \frac{1}{t} \int_0^t \frac{2as}{\frac{\pi}{2}\rho^2 + 2s\rho} \, \mathrm{d}s.$$

Equation (79) implies that $p_2^{wb} = 1 - \frac{2}{\lambda}(1 - \frac{2}{\lambda}) + O(a)$.

$\square$

4. Proofs of auxiliary results

This section contains the proofs of Lemmas 5, 7, 15, 19, 20 and 23. These lemmas are not concerned with multi-scale aspects.

We first explain the notation used in Lemma 5. Let $w \in C(\mathbb{R}^d)$, $w \geq 0$ be a weight. For a Radon-measure f we define

$$\|f\|_w := \sup_{\phi \in BC^0(\mathbb{R}^d), \|\phi\| \leq 1} \int |\phi(v) w(v) \, \mathrm{d}f(v)|.$$

Then $M_w = \{f \in (BC^0(\mathbb{R}^d)))^* | \ \|f\|_w < \infty\}$ is a Banach space of measures with norm $\|.\|_w$. To control convergence we introduce weighted spaces in time for X-valued functions, for some Banach space X

$$C^0_\rho([0,\infty), X) := \{u \in C^0([0,\infty), X) | \sup_{t\in[0,\infty)} (\exp(-\rho t)\|u(t)\|_X < \infty\}) \text{ with norm}$$
$$\|u\|_\rho := \sup_{t\in[0,\infty)} (\exp(-\rho t)\|u(t)\|_X) \, .$$

Proof of Lemma 5. The proof is based on a simple contraction argument. First, we note that $\|f_{t,k}\|_{(1+|v|)^2}$ is decreasing in t as $0 \leq L[f_{s,k-1}](v) < \infty$. Next we estimate $\exp(-\rho t)\|f_{t,k+1} - f_{t,k}\|_{1+|v|}$ for $0 \leq t < \infty$, with ρ chosen later. Let $\phi \in BC^0(\mathbb{R}^d)$ with $\|\phi\| \leq 1$, then consider

$$\exp(-\rho t) \left| \int_{\mathbb{R}^d} \phi(v)(1+|v|) \left(\mathrm{d}f_{t,k+1}(v) - \mathrm{d}f_{t,k}(v)\right) \right|$$
$$= \int_{\mathbb{R}^d} \phi(v)(1+|v|) \, \mathrm{d}f_0(v) \exp(-\rho t) \left| \exp\left(-\int_0^t L[f_{s,k}](v) \, \mathrm{d}s\right) - \exp\left(-\int_0^t L[f_{s,k-1}](v) \, \mathrm{d}s\right) \right|$$
$$\leq \int_{\mathbb{R}^d} \phi(v)(1+|v|) \, \mathrm{d}f_0(v) \exp(-\rho t) \int_0^t |L[f_{s,k}](v) - L[f_{s,k-1}](v)| \, \mathrm{d}s).$$

Because of the positivity of L, we obtain a Lipschitz constant of 1 for $\exp(-.)$. We have

$$\leq \int_{\mathbb{R}^d} \phi(v)(1+|v|) \, \mathrm{d}f_0(v) \kappa_d \left(\exp(-\rho t) \int_0^t \int_{\mathbb{R}^d} |\mathrm{d}f_{s,k}(v') - \mathrm{d}f_{s,k-1}(v')| \, |v - v'| \mathrm{d}s \right)$$

Then using the norms in $M_1(\mathbb{R}^d)$ and $M_{1+|v|}(\mathbb{R}^d)$ and splitting the exponential term, we obtain

$$\leq \int_{\mathbb{R}^d} \phi(v)(1+|v|) \, \mathrm{d}f_0(v) \kappa_d \Big(\int_0^t \exp(-\rho(t-s)) \Big[\exp(-\rho s)\|f_{s,k} - f_{s,k-1}\|_{1+|v|}$$
$$+ \exp(-\rho s)|v| \|f_{s,k} - f_{s,k-1}\|_1 \Big] \mathrm{d}s \Big)$$
$$\leq 2\kappa_d \int_{\mathbb{R}^d} \phi(v)(1+|v|)^2 \, \mathrm{d}f_0(v) \sup_{0\leq s<\infty} \left(\int_0^t \exp(-\rho(t-\tau)) \, \mathrm{d}\tau \right) \left(\exp(-\rho s)\|f_{s,k} - f_{s,k-1}\|_{1+|v|} \right)$$
$$\leq 2\kappa_d \|f_0\|_{(1+|v|)^2} \frac{1}{\rho} (1 - \exp(-\rho t)) \|f_k(.) - f_{k-1}(.)\|_\rho.$$

Thus for $\rho > 2\kappa_d \|f_0\|_{(1+|v|)^2}$ the sequence $(f_k)_{k\in\mathbb{N}}$ converges in $C^0_\rho([0,\infty), M_{1+|v|})$ by Banach's fixed point theorem and the limit solves $f_t = \exp(-\int_0^t L[f_s](v) \, \mathrm{d}s) f_0$. Hence f is differentiable and solves (11) for $t \in [0,\infty)$. Uniqueness of the solution of the integral equation also follows by the Banach fixed point theorem. On the other hand all solutions of (11) in $C^1([0,T], M_{1+|v|})$ have to satisfy the integrated form too, showing uniqueness of the solutions of (11). As $0 \leq f_t(v) \leq f_0(v)$, we also obtain $f_t \in M_{(1+|v|)^2}$. □

Proof of Lemma 7. We first show, that the implicit relation $\beta(i,t)$ in Theorem 2 is well-defined. For each particle it indicates whether it has undergone a collision: $\beta(i,t)$ jumps from 1 to 0 at the time of the collision. As the particles are removed after a collision, a collision can only occur when

$$\mathrm{dist}(z_i, z_{i'}, s) = a \text{ for some } i \neq i'.$$

This also takes multiple collisions into account, which lead to an undefined situation in hard-sphere collision dynamics, but as particles are removed here after a collision, the scattering state can be defined.
The distance $\mathrm{dist}(z_i, z_{i'}, s)$ is a continuous piece-wise affine function in s, except possibly a unique point, if there is an initial intersection, but then $\mathrm{dist}(z_i, z_{i'}, s) > a$ near this jump. There are only finitely many different pieces in a finite interval $[0, t]$, because $v(i) - v(i')$ is finite and only a finite number of coverings of the torus $\mathbb{T}^d$ can be visited in a finite time. Hence for every particle i, there are at most $N - 1$ possible collision times, i.e. the first time $\tau(i, i') \geq 0$ at which $\mathrm{dist}(z_i, z_{i'}, s) = a$ for each i'. The at most $N(N-1)/2$ possible times for collision of the particles $i = 1, \ldots, N$ can be well-ordered. So by inductively checking at all possible collision times $\tau(i, i')$, there exists a well-defined collision time for each particle i, at which it collides with an unscattered particle ($\beta(i, .)$ has a well-defined jump); or the particle remains unscattered itself for $[0, \infty)$ ($\beta(i)$ is constant), which shows the existence of $\beta(i, t)$.
To prove convergence of $\beta_k(i, t)$, defined in (16), to $\beta(i, t)$ as k tends to ∞, we first introduce some notation using the real scattering state $\beta(., .)$. Let τ_j be an ordering of the finite number of collision events described by $\beta(., .)$. The sets I_j are particles available for collision at time τ_j and C_j are those particle actually colliding. We define $I_1 = \{1, \ldots, N\}$ and $\tau_0 = 0$. For each $j \geq 1$ let $\tau_j > \tau_{j-1}$ and $C_j, I_{j+1} \subset I_j$ be recursively defined by

$$\begin{aligned}
&\min\{\mathrm{dist}(z_i, z_{i'}, \tau_j) \,|\, i \neq i' \in I_j\} = a \text{ for each } i \in C_j, \\
&\mathrm{dist}(z_i, z_{i'}, s) > a \text{ for all } i, i' \in I_j, s \in [\tau_{j-1}, \tau_j), \\
&I_{j+1} = I_j \setminus C_j.
\end{aligned}$$

It can be checked that $\beta(i, s) = 1$ if there exists $j \in \mathbb{N}$ such that $i \in I_j$ and $s \in [0, \tau_j]$. For all other choices of i and s we have that $\beta(i, s) = 0$. Clearly $\beta(i, \cdot)$ is constant within the intervals $(\tau_{j-1}, \tau_j]$. We will show using induction that for each $j \in \{1, 2, \ldots\}$ and each $k \geq j$

$$\beta_k(i, s) = \beta(i, s) \text{ if } s < \tau_j \text{ or } i \in I_1 \setminus I_j. \tag{84}$$

The claim is clear for $j = 1$. Assume now that the claim has been established up to j and let $k \geq j + 1$. We will show that

$$\beta_k(i, s) = \beta(i, s) \text{ if } s < \tau_{j+1} \text{ or } i \in I_1 \setminus I_{j+1}. \tag{85}$$

By the induction assumption (85) holds for $s \in [0, \tau_j]$ or $i \in I_1 \setminus I_j$ and we can assume from now that $s > \tau_j$.
Case 1. Let $i \in I_1 \setminus I_{j+1}$. We have to show that

$$\beta_k(i, s) = \beta(i, s) \text{ for all } k \geq j + 1. \tag{86}$$

Since $s > \tau_j$ we have that $\beta(i, s) = 0$. By (84) eq. (86) holds if $i \in I_1 \setminus I_j$, hence we can assume that $i \in C_j$. In this case there exists $i' \in C_j$ such that $\mathrm{dist}(z_i, z_{i'}, \tau_j) = a$. The induction assumption (84) implies that $\beta_{k-1}(i', \tau_{j-1}) = 1$ and thus

$$\mathrm{dist}(z_i, z_{i'}, \tau_j) = a\beta_{k-1}(i', \tau_j),$$

this implies by definition (16) that $\beta_k(i, s) = 0$.

Case 2. Let $i \in I_{j+1}$ and $s \in (\tau_j, \tau_{j+1})$. We show that

$$\beta_k(i,s) = \beta(i,s) = 1 \text{ for all } k \geq j+1. \tag{87}$$

Using definition (16) again, we have to ensure $\mathrm{dist}(z_i, z_{i'}, s) \geq a\beta_{k-1}(i', s)$ for all i' and $s \in (\tau_j, \tau_{j+1})$. If i' is such that $\beta_{k-1}(i', s') = 0$, then the condition holds trivially as always

$$\mathrm{dist}(z_i, z_{i'}, s') \geq 0 = a\beta_{k-1}(i', s').$$

So consider instead i' with $\beta_{k-1}(i', s') = 1$ for $s' \in (\tau_{j-1}, \tau_j)$. Then the induction assumption implies that $\beta_{k-1}(i', s') = \beta(i', s')$ and therefore $i' \in I_j$. Since $i \in I_{j+1}$ and $i' \in I_j$ it is not possible that $\mathrm{dist}(z_i, z_{i'}, \tau_j) = a$. Hence we obtain $\mathrm{dist}(z_i, z_{i'}, s) \geq a = a\beta_{k-1}(i', s)$ for $s \in (\tau_j, \tau_{j+1})$. This completes the induction step.
Since the number of particles is finite, eq. (18) is a consequence of (84). □

Proof of Lemma 19. To simplify the notation we define $\hat{\mathcal{C}} = \mathcal{C} \setminus \bar{\mathcal{C}}$. The assumption $A \subset \cup_{r=0}^{\infty} (\mathcal{C} \setminus \bar{\mathcal{C}})^r$ implies

$$\begin{aligned}
&\mathrm{Prob}_{\mathrm{ppp}}\left(\mathcal{C} \cap \omega \in A \mid \bar{\mathcal{C}} \cap \omega = \emptyset\right) = \mathrm{Prob}_{\mathrm{ppp}}\left(\hat{\mathcal{C}} \cap \omega \in A \mid \bar{\mathcal{C}} \cap \omega = \emptyset\right) \\
=&\mathrm{Prob}_{\mathrm{ppp}}\left(\hat{\mathcal{C}} \cap \omega \in A\right)
\end{aligned}$$

by independence since the sets $\hat{\mathcal{C}}$ and $\bar{\mathcal{C}}$ are disjoint. The last expression is an unconditional probability with respect to the Poisson-point process which can be evaluated explicitly using Definition 1:

$$\begin{aligned}
&\mathrm{Prob}_{\mathrm{ppp}}\left(\hat{\mathcal{C}} \cap \omega \in A\right) = \sum_{r=0}^{\infty} \mathrm{Prob}_{\mathrm{ppp}}\left(\hat{\mathcal{C}} \cap \omega \in A \cap \mathcal{C}^r\right) \\
=&\sum_{r=0}^{\infty} e^{-\mu(\hat{\mathcal{C}})} \frac{(\mu(\hat{\mathcal{C}}))^r}{r!} \times (\mu(\hat{\mathcal{C}}))^{-r} \int_{\hat{\mathcal{C}}^r} \mathrm{d}\mu^r(z) \chi_A(z) = e^{-\mu(\hat{\mathcal{C}})} \sum_{r=0}^{\infty} \frac{1}{r!} \int_{A \cap \mathcal{C}^r} \mathrm{d}\mu^r(z).
\end{aligned}$$

This proves formula (54). □

Proof of Lemma 15. This is a proof by induction over k. For $k = 1$, we immediately have in the tree description $\beta_1 = 1$ and $\beta_1(i,t) \equiv 1$ in the many particle dynamics. For $k > 1$, we consider a tree (m, ϕ) with root particle i^* which has r_1 particle on level two with subtrees $m_1, \ldots m_{r_1}$. For each subtree m_j with root particle j^* and time span s_{1j} we have $\beta_1(m_j) = \beta_{k-1}(j^*, s_{1j})$ by assumption. Then by (22) and the induction assumption

$$\beta_1(m) = \prod_{j=1\ldots r_l} (1 - \beta_1(m_j)) = \prod_{j=1\ldots r_l} (1 - \beta_{k-1}(j^*, s_{1j})). \tag{88}$$

Whereas considering (16), $\beta_k(i^*, s_1)$ can only be 1 if $\mathrm{dist}(z_{i^*}, z_j, s) \geq a\beta_{k-1}(j, s)$ for all s and $j \neq i^*$. Only the particles j^* with $j \in \{1, \ldots, r_l\}$ have some s with $\mathrm{dist}(z_{i^*}, z_{j^*}, s) \leq a$, namely s_{1j}. Hence we have $\beta_k(i^*, s_1) = 1$ if and only if $\beta_{k-1}(j^*, s_{1j}) = 0$ for all $j \in \{1, \ldots, r_l\}$. This is equivalent to the right hand side of (88) and hence $\beta_1(m)$ being 1, completing the proof. □

Proof of Lemma 20. We prove this by induction over k. For $k = 1$, this is just the definition in (46). Now assume that eq. (60) holds for $k - 1$. We split Ω in two parts. Firstly for $\Omega \subset \mathcal{G}(a) \cap \mathcal{MT}_{k-2}$, we obtain by (43) that $\hat{P}_k(\Omega) = \hat{P}_{k-1}(\Omega)$ and then the right-hand sides of (60) coincide as well.

For $\Omega \subset \mathcal{G}(a) \cap (\mathcal{MT}_k \setminus \mathcal{MT}_{k-2})$ a Borel set, we have by (45) and (59)

$$\hat{P}_k(\Omega) = \int_{\mathbb{T}^d\times\mathbb{R}^d} \mathrm{d}P_1(\phi_1(Z_1))\, e^{-\hat{\Gamma}_1(\Phi_1(Z_1))} \int_{(\mathbb{T}^d\times\mathbb{R}^d)^{r_2}} \mu^{r_2}(\Phi_2(Z_1Z_2))$$
$$\dots e^{-\hat{\Gamma}_{k-1}(\Phi_{k-1}(Z_1\dots Z_{k-1}))} \int_{A_k(\Omega,\Phi_{k-1}(Z_1\dots Z_{k-1}))} \mathrm{d}\mu^{r_k}(Z_1Z_2\dots Z_k)$$
$$= \sum_{m\in\mathcal{T}_k\setminus\mathcal{T}_{k-2}} \int_{A(\Omega)} \mathrm{d}\mu^{\#m}(z)\, e^{-\sum_{j<k}\hat{\Gamma}_j(\Phi(z))},$$

where we observe that $A_k(\Omega, \Phi_{k-1}(Z_1 \dots Z_{k-1}))$ is empty for $m \in \mathcal{T}_{k-1} \setminus \mathcal{T}_{k-2}$. This empty integral is then evaluated as 0 and we obtain (60) for all $\Omega \subset \mathcal{G}(a) \cap \mathcal{MT}_k$. $\square$

Proof of Lemma 23. For each $\xi \in \mathbb{R}^d \setminus \{0\}$ we define the cone

$$M(\xi, a) = \left\{v \in \mathbb{R}^d \,|\, (v\cdot\xi)^2 \geq (|\xi|^2 - a^2)|v|^2\right\}.$$

We first observe that $M(\xi, a) \subset M(\frac{1}{|\xi|}\xi, a)$ for all $\xi \in \mathbb{R}^d$ with $|\xi| \geq 1$, i.e. in particular for $\xi \in \mathbb{Z}^d\setminus\{0\}$. Letting $c(a) := \sup\{\int_{M(\xi,a)} \mathrm{d}f_0(v) \,|\, \xi \in \mathbb{Z}^d\setminus\{0\}\}$, we see, if $\limsup_{a\to 0} c(a) > 0$, then by normalizing there exists a converging sequence of directions ξ_j in S^{d-1} such that $\xi_j \to \xi$ and $\int_{M(\xi_j,a_j)} \mathrm{d}f_0(v) > c$. This implies with $M(\xi_j, a_j) \subset M(\xi, a_j + |\xi - \xi_j| + 2\sqrt{|\xi - \xi_j|})$, that $\limsup_{a\to 0} \int_{M(\xi,a)} \mathrm{d}f_0(v) > c$ in contradiction to assumption (6). Hence we have that $c(a) = o(1)$ as $a \to 0$. For each $v \in R(t, a)$ such that $|v| \leq \bar{V}$ there exists $\xi(v) \in \mathbb{Z}^d\setminus\{0\}$ such that $|\xi(v)| \leq \bar{V}t + a$ and $v \in M(\xi(v), a)$, i.e. each velocity $v \in R(t, a)$ is an element of one of at most $(2t\bar{V} + 2a)^d$ cones. Thus we obtain, using (5),

$$\int_{R(t,a)} (1 + |v'|)\, \mathrm{d}f_0(v') \leq \int_{R(t,a)\cap\{|v'|\leq\bar{V}\}} (1 + |v'|)\, \mathrm{d}f_0(v') + \int_{\{|v'|>\bar{V}\}} (1 + |v'|)\, \mathrm{d}f_0(v')$$
$$\leq (1 + \bar{V})(2t\bar{V} + 2a)^d c(a) + K_{\mathrm{ini}}/\bar{V},$$

with $K_{\mathrm{ini}} = \int_{\mathbb{R}^d} \mathrm{d}f_0(v)\,(1+|v|)^2$. So choosing first $\bar{V}$ large the second term is small. Then choose a so that the first term is small, which completes the proof of the equation (66). $\square$

5. Discussion

In this paper we propose and develop a new method that allows us to derive and justify effective continuum limits as scaling limits of large interacting particle systems. We consider the conceptually simplest situation of kinetic annihilation where each particle moves with constant velocity until it interacts with another particle. After the collision the collided particles are removed from the system. The transport term could be dropped by considering spatially homogenous initial data. The analysis of kinetic annihilation with transport term will be the subject of a forthcoming paper.

It would be highly desirable to generalize our approach so that also the case of collisional dynamics can be treated. The main difficulty arises from the fact that although the concept of the collision trees can be adapted it is harder to obtain lower bounds on probabilities of good events. The reason is that in the case of collisional dynamics the trees consist of two different types of nodes:

1. Destructive collisions which prevent the root particle from being in a certain state. These collisions correspond to the loss term in the Boltzmann equation.
2. Constructive collisions which explain the momentum changes of observed particles. These collisions correspond to the gain term in the Boltzmann equation and do not occur in the gainless case.

Formally, the likelihood of trees with constructive and destructive nodes can be computed with a formula analogous to (24), but due to the presence of two different types of nodes the integrand changes its sign. Moreover, it can be checked that for sufficiently large t, the integrand is not absolutely integrable, i.e. the integral only makes sense when cancelation effects are taken into account. These cancelation effects are the probabilistic analogue of the fact that solutions of the homogeneous Boltzmann equation

$$\frac{\partial f}{\partial t} = Q_+[f,f] + Q_-[f,f]$$

only exist globally in time due to cancellation effects, in the sense that the lossless Boltzmann equation $\frac{\partial f}{\partial t} = Q_+[f,f]$ does not admit global solutions, see [IS87]. For this reason it is currently unclear, whether almost sharp lower bounds on the likelihood of good trees can be obtained in this way.

Appendix A. Notation

Symbol	Meaning
a	diameter of the balls
N	number of particles
$n = a^{1-d}$	intensity of the Poisson measure for the initial positions on $\mathbb{T}^d$
(u, v)	phase space variables in $\mathbb{T}^d \times \mathbb{R}^d$
f_0	initial velocity distribution, element of $PM(\mathbb{R}^d)$
$\mathrm{Prob}_{\mathrm{tppp}}$	probability of the Poisson-point process of the initial data, Definition 1
f_k	approximate solution of the gainless, homogeneous Boltzmann equation, (15)
$PM(\mathbb{R}^d)$	probability measure on $\mathbb{R}^d$
$\mathcal{H}^d$	d-dimensional Hausdorff measure
$M_+(\mathbb{T}^d \times \mathbb{R}^d)$	non-negative measures on $\mathbb{T}^d \times \mathbb{R}^d$
$M_w(\mathbb{R}^d)$	measures with weight function w, after Lemma 5
$\beta^{(a)}(i, t)$	scattering state ($= 1$ unscattered, $= 0$ scattered) of particle i at time t, (16)
$\beta_k(i, t)$	scattering state when restricting to tree of height k, see (16,22) and Lemma 15
$\mathcal{T}$	$\subset \cup_{i=1}^{\infty} \mathbb{N}^i$ set of tree skeletons, Definition 8
m	$\in \mathcal{T}$ tree (skeleton), Definition 8
l	$\in m$ a node in a tree, Definition 8
$\bar{l}$	the parent of node l, Definition 8
$\lvert l \rvert$	height of a node ($= i$ if $l \in \mathbb{N}^i$), Definition 8
r_l	number of children of node l, (21)
(u_l, v_l, s_l, ν_l)	$\in \mathbb{T}^d \times \mathbb{R}^d \times [0, \infty) \times S^{d-1}$ data on node l with u_l, v_l initial data, ν_l collision parameter and s_l collision time, Definition 8
$\mathcal{MT}$	marked trees with collision data, Definition 8
$\mathcal{E}(m)$	$\subset \mathcal{MT}$ trees with skeleton m, (19)
$\Phi = (m, \phi)$	$\in \mathcal{MT}$ tree (with collision data)
$P_{t,k}$	idealized probability, (24)
$P_{\underline{t},1}$	distribution of root, (46)
$\mathrm{d}\lambda_l$	simplified idealized distribution at node l, (29)
$\hat{P}_{t,k}$	empirical distribution, (42)
$R(t, a)$	$\subset \mathbb{R}^d$ resonant initial velocities, (47)
$\mathcal{G}(a)$	$\subset \mathcal{MT}$ good trees, Definition 18
$\hat{\mathcal{G}}(a_0)$	$\subset \mathcal{G}(a_0)$ good trees with additional desirable properties, (67)
γ_l	integrated collision rate of particle l (idealized), (26)
$\Gamma(j)$	joint integrated collision rate of particles of height j (idealized), (25)
$\hat{\Gamma}(j)$	joint integrated collision rate of particle of height j (empiric), (57)
$C_l = C_l(\phi)$	colliding initial values of particle at node l, Definition 18
$\mathcal{C}(k)$	$:= \bigcup_{l \in m \cap \mathbb{N}^k} C_l \subset \mathbb{T}^d \times \mathbb{R}^d$, (55)
$\bar{\mathcal{C}}(k)$	$:= \bigcup_{\lvert l \rvert < k} C_l \subset \mathbb{T}^d \times \mathbb{R}^d$, after eq. (55)
$\hat{\mathcal{C}}(k)$	$:= \mathcal{C}(k) \setminus \bar{\mathcal{C}}(k)$, (58)
BC^0	bounded continuous functions

Acknowledgments

The authors are grateful to Davide Marenduzzo for careful numerical simulations. The authors are very thankful for the detailed comments by a referee, which improved considerably the presentation of the paper.

References

[BBS83] C. Boldrighini, L.A. Bunimovich, Y.G. Sinai. On the Boltzmann equation for the Lorentz gas. *J. Stat. Phys.* **32** (1983), 477–501.

[CIP94] C. Cercignani, R. Illner, M. Pulvirenti. *The Mathematical Theory of Dilute Gases.* Applied Mathematical Sciences, Vol **106**, Springer Verlag (1994).

[CDPTW03] F. Coppex, M. Droz, J. Piasecki, E. Trizac, P. Wittwer. Some exact results for Boltzmann's annihilation dynamics. *Phys. Rev. E* **79** (2003), 21103.

[DL90] R. DiPerna, P.L. Lions. On the Cauchy Problem for Boltzmann Equations: Global Existence and Weak Stability, *Ann. Math.* **130** (1989), 321–366.

[DFPR95] M. Droz, L. Frachebourg, J. Piasecki, P.-A. Rey. Ballistic annihilation kinetics for a multi-velocity one-dimensional ideal gas. *Phys. Rev. E* **51** (1995), 5541–5548.

[Dur] R. Durrett, *Probability: Theory and Examples.* 3^{rd} ed. Duxbury (2004).

[EF85] Y. Elskens, H. Frisch. Annihilation kinetics in the one-dimensional ideal gas. *Phys. Rev. A* **31** (1985), 3812–3816.

[Gal70] G. Gallavotti. Rigorous theory of Boltzmann equation in the Lorentz gas, preprint *Nota interna* **358**, Univ. di Roma (1970).

[Hil00] D. Hilbert, Mathematical problems. Reprinted from Bull. Amer. Math. Soc. 8 (1902), 437–479. *Bull. Amer. Math. Soc. (N.S.)* **37** (2000), 407–436.

[IP89] R. Illner, M. Pulvirenti. Global validity of the Boltzmann equation for two- and three-dimensional gas in vacuum. Erratum and improved result. *Comm. Math. Phys* **121** (1989), 143–146.

[IS87] R. Illner, M. Shinbrot. Blow-up of solutions of the gain-term only Boltzmann equation. *Math. Meth. in the Appl. Sci.* **9** (1987), 251–259.

[Kal05] O. Kallenberg. *Probabilistic Symmetries and Invariance Principles*, Probability and its Applications, Springer Verlag (2005).

[KS88] J. Krug, H. Spohn. Universality classes for deterministic surface growth. *Phys. Rev. A* **38** (1988), 4271–4283.

[Lan75] O. Lanford. Time evolution of large classical systems. In: Dynamical systems, theory and applications. ed. by J. Moser, Lecture notes in physics, **38**, Springer Verlag, pp. 1–111. (1975).

[LN80] R. Lang, X. Nguyen. Smoluchowski's theory of coagulation holds rigorously in the Boltzmann-Grad limit. *Z. Wahrs. Verw. Geb.* **54** (1980), 227–280.

[MT08] K. Matthies, F. Theil. Validity and non-validity of propagation of chaos. In: *Analysis and stochastics of growth processes* ed. by P. Mörters et. al., Oxford University Press, pp. 101–119 (2008).

[Pia95] J. Piasecki. Ballistic annihilation in a one-dimensional fluid. *Phys. Rev. E* **51** (1995), 5535–5540.

[PTD02] J. Piasecki, E. Trizac, M. Droz. Dynamics of ballistic annihilation. *Phys. Rev. E* **65** (2002), 66111.

[Spo78] H. Spohn. The Lorentz process converges to a random flight process. *Comm. Math. Phys.* **60** (1978) 277–290.

[Spo91] H. Spohn. *Large scale dynamics of interacting particles.* Texts and Monographs in Physics, Springer Verlag (1991).

[Sz91] A. Sznitman. Topics in the propagation of chaos. In P. Hennequin (Ed.), *Ecole d'Eté de Probabilités de Saint-Flour 1989*, Lecture Notes in Mathematics, **1464** , pp. 165–251 (1991).

[Ush88] K. Ushiyama. On the Boltzmann-Grad limit for the Broadwell model of the Boltzmann equation. *J. Stat. Phys.* **52** (1988), 331–355.

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom
E-mail address: K.Matthies@maths.bath.ac.uk

Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail address: theil@maths.warwick.ac.uk