

A SEMIGROUP APPROACH TO THE JUSTIFICATION OF KINETIC THEORY*

KARSTEN MATTHIES[†] AND FLORIAN THEIL[‡]

Abstract. This paper develops a method to rigorously show the validity of continuum description for the deterministic dynamics of many interacting particles with random initial data. We consider a hard sphere flow where particles are removed after the first collision. A fixed number of particles is drawn randomly according to an initial density $f_0(u, v)$ depending on d -dimensional position u and velocity v . In the Boltzmann–Grad scaling, we derive the validity of a Boltzmann equation without gain term for arbitrary long times, when we assume finiteness of moments up to order two and initial data that are L^∞ in space. We characterize the many-particle flow by collision trees which encode possible collisions. The convergence of the many-particle dynamics to the Boltzmann dynamics is achieved via the convergence of associated probability measures on collision trees. These probability measures satisfy nonlinear Kolmogorov equations, which are shown to be well-posed by semigroup methods.

Key words. Boltzmann equation, Boltzmann–Grad limit, validity, kinetic annihilation, deterministic dynamics, random initial data, semigroups, Kolmogorov equation

AMS subject classifications. 82C40, 35Q20, 37L05, 82C22, 76P05, 60K35

DOI. 10.1137/120865598

1. Introduction. Deriving continuum models as a scaling limit of atomistic particle dynamics is a fundamental problem of mathematical physics. The aim is to prove the validity of continuum equations like the Boltzmann equation to describe the effective behavior of many particle dynamics. The first rigorous derivation was given by Lanford [Lan75] for short times using the Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy. The problem of convergence of this hierarchy was partially overcome by using sufficiently small initial data on unbounded domains [IP89] or by considering linear variants; related results can be found in [Spo78, BBS83, Spo91, CIP94].

In this paper we consider kinetic annihilation, a 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 of n particles in the phase space $U \times \mathbb{R}^d$ ($U = \mathbb{T}^d$ is the unit torus) are drawn independently with some density $f_0 \in L^1(U \times \mathbb{R}^d)$. 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. Kinetic (or ballistic) annihilation has been studied extensively in the physics literature (see [CDPTW03, Pia95, DFPR95, PTD02]), including a proof that the Boltzmann approximation does not hold in one space dimension [EF85]. The model can be used, e.g., to model growth and coarsening of surfaces; see [KS88]. A closely related system is given by coagulating Brownian particles where the continuum limit is given by a system of reaction diffusion equation; see [LN80, Szn87]. In this case

*Received by the editors February 10, 2012; accepted for publication (in revised form) October 17, 2012; published electronically December 18, 2012.

<http://www.siam.org/journals/sima/44-6/86559.html>

[†]Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom (k.matthies@maths.bath.ac.uk).

[‡]Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom (theil@maths.warwick.ac.uk).

the validity of the continuum equation is known for large times. In [Rez04, HR07] stochastic variants of the hard ball evolution and coagulating Brownian particles are studied where the collisions between the individual particles are random in the sense that two particles at distance r collide at any given moment with a rate $V(r)$. These systems have better ergodic properties, and the justification of the continuum limit does not require the tracking of pair distributions.

There are two slightly different strategies for justifying the scaling limit for deterministic collisions. In [Lan75] the focus is on k -particle projections of the many-body evolution. This leads to involved expressions for the correlations which are eventually controlled via the BBGKY hierarchy. In this paper we follow a strategy which is similar to the approach in [Szn87]. In addition to the k -particle marginal, information about the past evolution is kept as well, and this leads to transparent expressions for the correlations. A preliminary version was first introduced in [MT08, MT10].

In this article we combine our approach with semigroup theory. This allows us to consider spatially heterogeneous initial distributions. The challenge is that we have to include the transport term in our analysis, which leads to more stringent regularity requirements for the initial distribution.

We consider the evolution of n balls of diameter a and with position $u(i, t) \in U \subset \mathbb{R}^d$ for $i \in \{1, \dots, n\}$ with $d \geq 2$ and respective velocity $v(i, t) \in \mathbb{R}^d$. Our main interest is the kinetic limit, when the number of particles n tends to infinity and the initial values $(u(i, 0), v(i, 0))$ are independent identically distributed (iid) random variables distributed according to some initial distribution f_0 . The diameter a of the particles is coupled to the number n by the Boltzmann–Grad scaling, which is in the easiest form

$$(1.1) \quad na^{d-1} = 1.$$

The final aim is to analyze the situation, where the particles interact via some suitable potential, like a hard sphere one. We are going to compare deterministic continuum descriptions with the empirical density given for all open sets $A \subset U \times \mathbb{R}^d$ and any fixed time by the number of particles in A divided by the total number of particles. Our main result is that the density f of the continuum description solves the nonlinear Boltzmann equation

$$(1.2) \quad \partial_t f + v \cdot \nabla_u f = -Q_-[f, f],$$

where $Q_- \geq 0$ is the collision operator accounting for the losses. The collision operator can be easily derived for hard core potentials in a situation of completely independent particles with density $f(u, v, t)$. Particles with velocities v and v' collide at position u with a given probability depending on v and v' and impact parameter $\nu \in S^{d-1}$, which encodes the collision angle of two particles (see Figure 1). In the density there is a loss at (u, v) and (u, v') . The loss operator has the form

$$Q_-[f, f](u, v) = \int_{\mathbb{R}^d} \int_{S^{d-1}} f(u, v') f(u, v) [(v - v') \cdot \nu]_+ d\nu dv'.$$

In the case of collisional dynamics the loss is balanced by the corresponding gain at $(u, v_*) = (u, v - ((v - v') \cdot \nu)\nu)$ and $(u, v'_*) = (u, v' + ((v - v') \cdot \nu)\nu)$ with the consequence that the Boltzmann equation is augmented by the appropriate gain term.

In [MT10] we analyze a situation of spatially homogeneous initial data, which corresponds to a version of (1.2), where the transport term $v \cdot \nabla_u f$ vanishes. In the present paper we will allow for heterogeneous initial data reintroducing the transport

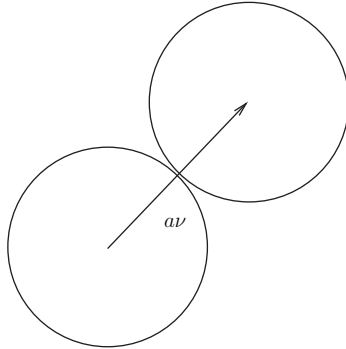


FIG. 1. The impact parameter v .

term. To handle the transport term we will restrict our attention to initial densities f_0 which are absolutely continuous with respect to the Lebesgue measure,

$$(1.3) \quad f_0 \in L^1(U \times \mathbb{R}^d) \text{ with } f_0 \geq 0 \text{ and } \int_{U \times \mathbb{R}^d} f_0(u, v) \, du \, dv = 1,$$

and have finite total energy:

$$(1.4) \quad \int_{U \times \mathbb{R}^d} (1 + |v|)^2 f_0(u, v) \, du \, dv = K < \infty.$$

We require the u -marginal of f_0 to be in $L^d(U)$ to ensure that particles overlap at any given point only with probability zero (see section 7) and that the energy density and its transported versions are also bounded in $L^\infty(U)$,

$$(1.5) \quad K_\infty = \text{ess sup}_{(t,u) \in (0,T) \times U} \int_{\mathbb{R}^d} (1 + |v|^2) f_0(u - tv, v) \, dv < \infty.$$

Note that (1.5) and the continuity of the L^∞ norm with respect to mollification implies

$$(1.6) \quad \text{ess sup}_{u \in U} \int_{\mathbb{R}^d} (1 + |v|^2) f_0(u, v) \, dv \leq K_\infty < \infty.$$

The main result (Theorem 2.1) is a rigorous justification of (1.2) if f_0 fulfills (1.3), (1.4), and (1.5). A key element in the proof is an intermediate layer of description between the complicated n -body evolution and the one-body distribution $f(\cdot, \cdot, t)$. This layer consists of trees which describe the history of collisions of an individual particle and its potential scattering particles. This extra layer allows on one hand a relatively easy description of the limiting (idealized) distribution P_t ; see the definition in (4.39). On the other hand we can estimate the error between the empirical distribution \hat{P}_t created by the n -body evolution and the idealized distribution P_t ; see Proposition 5.5.

In contrast to [MT10], where we used explicit formulas for the distributions, we derive nonlinear Kolmogorov equations for the evolution of the probability measures P_t and \hat{P}_t with time t . As we are essentially describing the evolution of low-dimensional marginals, the Kolmogorov equations are quadratic in the measure. A key result is the derivation of the Kolmogorov equation for \hat{P}_t which accounts for the correlations caused by the history of the evolution.

By fixing one of the arguments of the Kolmogorov operator it will be possible to apply general semigroup theory to the idealized evolution. A fixed-point argument then provides the existence of a nonlinear semigroup. The desired convergence of the multibody empirical distribution to the idealized one in the Boltzmann–Grad limit then follows with relative ease. The final step is to derive the density description $f_t(\cdot, \cdot)$ as a marginal from the distribution of trees. The Boltzmann equation will then appear naturally from the differential equation for the distribution on trees.

Allowing heterogenous initial data requires a number of additions to the methods in [MT10], because several new error terms are created by the spatial heterogeneity. A careful analysis of regularity of the initial data (1.3) f_0 is needed to deal with concentration phenomena in position space and to obtain solutions to (1.2). We will consider here a bounded domain with periodic boundary data, i.e., $U = \mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$.

On the level of partial differential equations several formulations of (1.2) are relevant. In a particular form this equation is well-defined for L^1 -data with respect to the space coordinate u . We require higher spatial regularity in the derivation, such that we can obtain standard mild and weak solutions of (1.2). Following ideas in [MT10], L^2 regularity is enough to prove tightness of the self-similar tree measure by deriving bounds on the expected number of nodes in the trees. We need L^d to obtain good bounds on the initial overlap of particles. In the current paper we impose L^∞ assumptions on the spatial energy density for simplicity of presentation.

It is noteworthy that the well-posedness of the Boltzmann equation in some function space does not imply that the limit of the single-particle distribution is a solution of the Boltzmann equation; an explicit counterexample has been constructed in [MT10].

In the current paper we prove all required regularity for finite times through a simple a priori bound of the solutions of (1.2) due to the sign of the right-hand side. However, we expect that with growing complexity more involved estimates will be required. The analytical understanding of various aspects of kinetic equations has progressed significantly within the last 25 years. A crucial tool is the gain of regularity and compactness through velocity averaging lemmas for various equations [Ag84, GLPS99, Ge90, GG92, DP01, DLM91]; for further references see the review [Per04]. The existence of renormalized solutions to the full Boltzmann equation [DL89a] uses transport theory as in [DL89b]. These tools are also used in [Rez04] and [BGL93, LM01a, LM01b, GS04] with the aim to derive the incompressible Navier–Stokes equation through scaling of solutions of the Boltzmann equation.

The paper is organized as follows. In section 2 we will describe the setup and formulate the main result. In section 3 the collision trees are introduced. Various probability distributions are considered in sections 4 and 5. The main theorem will be proved in section 6, by deriving the effective single particle dynamics. In section 7 we discuss spatial concentrations.

2. Setup and main result. We define the multibody evolution in the following way. Let $f_0 \in L^1(U \times \mathbb{R}^d)$ be a density of initial conditions. For each $n \in \mathbb{N}$ consider the random variable

$$(2.1) \quad (u^0, v^0) = (z_1, \dots, z_n) \in (U \times \mathbb{R}^d)^n$$

with z_1, \dots, z_n iid according to f_0 and a determined by (1.1) giving a probability measure Prob_a . The particles evolve by force-free Newtonian dynamics with initial conditions $(u_i^0, v_i^0) \in U \times \mathbb{R}^d$ for $i = 1, \dots, n$,

$$(2.2) \quad u_i(t=0) = u_i^0, \quad v_i(t=0) = v_i^0,$$

according to the differential equations

$$(2.3) \quad \begin{aligned} \dot{u}_i(t) &= v_i(t), \\ \dot{v}_i(t) &= 0. \end{aligned}$$

The scattering state (1 for unscattered, 0 for scattered and removed) for each particle $i = 1, \dots, n$ and time t is defined by

$$(2.4) \quad \beta_i^{(a)}(t) = \begin{cases} 1 & \text{if } d(z_i, z_{i'}, s) > a\beta_{i'}^{(a)}(s) \text{ and } |u_i^0 - u_{i'}^0| > a \text{ for all } s \in [0, t], i' \neq i, \\ 0 & \text{else,} \end{cases}$$

where a will depend on n and where the distance of particles on the torus U with data $z_i = (u_i, v_i)$ and $z_{i'} = (u_{i'}, v_{i'})$ is

$$d(z_i, z_{i'}, s) = |u_i(s) - u_{i'}(s)|_U = |u_i^0 - u_{i'}^0 + s(v_i^0 - v_{i'}^0)|_U.$$

This means in particular that particles are removed if they overlap at time $t = 0$. See [MT10] for a proof that $\beta_i^{(a)}(t)$ is well-defined. We compare the multibody evolution with the single-body description $f : U \times \mathbb{R}^d \times [0, \infty) \rightarrow \mathbb{R}$,

$$(2.5) \quad \begin{aligned} \partial_t f + v \cdot \nabla_u f &= -Q_-[f, f], \\ f(u, v, 0) &= f_0(u, v), \end{aligned}$$

where

$$(2.6) \quad Q_-[f, g](u, v) = L[g](u, v) f(u, v) = \left(\int_{\mathbb{R}^d} g(u, v') \kappa_d |v - v'| dv' \right) f(u, v)$$

is the loss term and κ_d is the volume of the $(d - 1)$ dimensional unit ball, in particular, $\kappa_2 = 2$ and $\kappa_3 = \pi$. We will consider mild solutions of (2.5), which are functions $f \in C^0([0, T], L^1(U \times \mathbb{R}^d))$ with

$$(2.7) \quad f_t = S_t f_0 - \int_0^t S_{t-s} Q_-[f_s, f_s] ds$$

in $L^1(U \times \mathbb{R}^d)$ for all $t \in [0, T]$, where S_t is the strongly continuous linear semigroup given by $S_t h(u, v) = h(u - tv, v)$.

THEOREM 2.1. *Let $f_0 \in L^1(U \times \mathbb{R}^d)$ with $d \geq 2$ be an initial distribution fulfilling (1.3), (1.4), (1.5). For $n \in \mathbb{N}$, consider the evolution of (2.3) with initial conditions (2.2) as in (2.1). The diameter a is coupled to n via the Boltzmann–Grad scaling*

$$(2.8) \quad na^{d-1} = 1.$$

Then the density of the unscattered particles converges to a solution of the Boltzmann equation in the sense that for all $\varepsilon > 0$ and all open $A \subset U \times \mathbb{R}^d$ uniformly for t in a compact set

$$(2.9) \quad \lim_{a \rightarrow 0} \text{Prob}_a \left(\left| \frac{1}{n} \# \left\{ i \mid (u_i(t), v_i(t)) \in A, \beta_i^{(a)}(t) = 1 \right\} - \int_A f_t(u, v) du dv \right| > \varepsilon \right) = 0,$$

where $f_t(\cdot, \cdot) = f(\cdot, \cdot, t)$ is the unique mild solution of (2.5). Furthermore, there exists a sequence $a_k \rightarrow 0$ and corresponding particle numbers n_k , such that with probability 1

$$(2.10) \quad \frac{1}{n_k} \sum_{i=1}^{n_k} \beta_i^{(a_k)}(t) \delta(\cdot - (u_i(t), v_i(t))) \xrightarrow{*} f_t$$

weak-* in $M(U \times \mathbb{R}^d)$ (the space of unsigned Radon measures) as $k \rightarrow \infty$ with δ denoting the Dirac distribution.

Remark 2.2.

- (i) The number of particles n is fixed for given diameter a , unlike in [MT10], where it was a random number given by a Poisson distribution with intensity a^{1-d} . So here we consider a canonical ensemble as opposed to a grand canonical ensemble in the easier case. As we need some control of correlations to prove convergence in probability (2.9), proofs would not be much easier for a grand canonical ensemble.
- (ii) For other notions of solutions of (2.5) see Proposition 4.10.
- (iii) Some effects of spatial concentration are analyzed in section 7, and concentration effects in velocity are ruled out via the absolute continuity with respect to the Lebesgue measure.
- (iv) A larger class of initial distributions like $f_0 \in L^d(U, BC^0(\mathbb{R}^d)^*)$ with some additional nonconcentration assumptions in velocity space seems to be conceivable but is not considered for presentational reasons.
- (v) The convergence of k -particle distribution functions to a product of f_t can be shown for every fixed k using the same method. This gives a connection to the classical derivation for short times using the BBGKY hierarchy, which was applied to the simpler problem of coagulation by Lang and Nguyen [LN80]. 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, a rigorous justification of the corresponding Boltzmann equation was obtained by restarting the procedure at small positive time. The BBGKY hierarchy could also be applied to the ballistic annihilation model, but this would require bounds on exponential moments.

3. Collision trees. We introduce the intermediate layer of collision trees to analyze the multibody dynamics. Each node of a collision tree corresponds to a particle. All nodes l except the root node are marked with information encoding a collision with another particle which corresponds to one of the tree nodes. The root node is marked with the initial position and velocity (u, v) of the corresponding particle.

Collisions happen in the gainless case, considered here, if $|u + sv - (u' + sv')|_U \leq a$ for some time $s \in [0, t]$ and some $(u', v') = (u_i^0, v_i^0)$ for some particle i . Given the time t and the set of all initial states, the tree of particle 1 is defined recursively. The children of the root node correspond to particles which intersect the path of the root particle up to time t . The child nodes are marked with the velocity v_l , collision time s_l , and impact parameter $\nu_l = \frac{1}{a}(u + s_l v - (u_l + s_l v_l))$. This rule is applied recursively to every child node with t replaced by s_l .

After this preparation it is easy to see that the scattering state of the root node is equal to 1 if and only if either the tree has just one node or the scattering states of

each child node is equal to 0. Thanks to the recursive definition of the tree and the fact that the scattering state of each leaf is 1, the scattering state of the root particle is thus a simple function of the tree structure.

Due to the finiteness of the number of particles and number of possible collisions in finite time for given velocities, the trees have finite size. We will later show that the size of the trees relevant in the description of (2.3) is uniformly bounded as $n \rightarrow \infty$ in our scaling. To compare the dynamics of several particles, we will consider “trees” with α roots, which is a forest in graph theory language. The number α is fixed; in particular, the behavior with $\alpha = n$ will not be considered.

We use the following notation for marked trees. We start with the standard graph theoretic notion of a rooted tree, i.e., an acyclic graph with a tagged vertex denoted as root. We use a partial order $>_p$ on trees. We say $k >_p l$ if k is on the unique simple path which connects l to the root. Note that this is opposite to the standard graph theoretic order, but it is more suitable in our context. We will denote by \bar{l} the first node on the simple path from a node l to the root and call \bar{l} the parent of l , whereas l is a child of \bar{l} .

We will use the notation (m, E) for a rooted tree, where m is the set of nodes and E the set of edges. The set of rooted trees will be denoted by \mathcal{T} .

DEFINITION 3.1 (marked trees). *Let $Y = \mathbb{R}^d \times [0, \infty) \times S^{d-1}$ be the set of child markers and $Y^* = U \times \mathbb{R}^d$ be the set of root markers. The collision trees \mathcal{MT} is the set of mappings from trees to $Y \cup Y^*$ such that the collision times respect the partial order of the vertices:*

$$\mathcal{MT} = \{((m, E), (u, v), (v_l, s_l, \nu_l)_{l \in m \setminus \text{root}}) : (m, E) \in \mathcal{T} \text{ and } s_l < s_k \text{ if } l <_p k\}.$$

The markers induce a finer partial order “ $<$ ” on the set of vertices:

$$l < k \text{ if there exists } l' \geq_p l, k' \geq_p k \text{ such that } \bar{l}' = \bar{k}' \text{ and } s_{l'} < s_{k'}.$$

The distance between two trees Φ and Ψ is defined as

$$d(\Phi, \Psi) = \begin{cases} \min \left\{ 1, \max_{l \in m(\Phi)} |\Phi_l - \Psi_l|_\infty \right\} & \text{if } (m, E)(\Phi) = (m, E)(\Psi), \\ 1 & \text{else.} \end{cases}$$

By $\tau(\Phi)$ we denote the final collision time

$$\tau(\Phi) = \begin{cases} \max \{s_l : \bar{l} = \text{root}\} & \text{if } \#m(\Phi) > 1, \\ 0 & \text{else,} \end{cases}$$

and

$$\mathcal{MT}_t = \{\Phi \in \mathcal{MT} : \tau(\Phi) = t\}$$

is the set of trees where the final collision takes place at t . For each node $l \in m(\Phi)$ the initial position $u_l \in \mathbb{T}^d$ is computed via the recursive formula

$$(3.1) \quad u_l = \begin{cases} u & \text{if } l = \text{root,} \\ u_{\bar{l}} + s_l(v_{\bar{l}} - v_l) + av_l & \text{if } \bar{l} \text{ is the parent of } l. \end{cases}$$

We will often write $u(\Phi)$ and $v(\Phi)$ instead of $u_{\text{root}}(\Phi)$ and $v_{\text{root}}(\Phi)$ as well as τ instead of $\tau(\Phi)$.

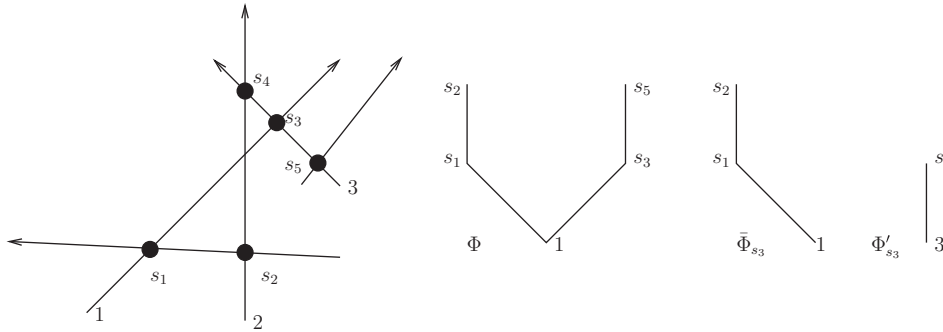


FIG. 2. Initial positions and velocities of five particles. The bullets indicate the positions where the particles are potentially scattered. The collision at time s_4 can be ignored in the left-hand tree Φ with root particle 1, as $s_4 > s_3$ and $s_4 > s_2$. The middle tree is the pruned tree $\bar{\Phi}$ with root 1. The right tree is the extracted subtree Φ' , which is obtained by using the colliding particle 3 (at time s_3) as the new root.

We will also consider trees generated by several particles, in this case $\Phi \in \mathcal{MT}^\alpha$, $\alpha \in \{1, 2, \dots\}$. The set $Y = \mathbb{R}^d \times [0, \infty) \times S^{d-1}$ denotes initial velocity v , collision time s , and impact parameter ν . The root marker $Y^* = U \times \mathbb{R}^d$ characterizes the initial position and velocity of the root particle. Some examples of collision trees are given in [MT10, Figure 1,2].

The definition of the evolution of the set of trees is based on two elementary operations: extraction of subtrees and pruning. We consider subtrees Φ' , where the new root corresponds to the particle which creates the final collision l with the root of Φ . The subtree Φ' contains all child nodes of the final collision and recursively all of their children, etc., that is, all nodes k with $l \geq_p k$. The pruned tree $\bar{\Phi}$ is obtained by removing all nodes of Φ' (and respective edges) from Φ ; for an illustration see Figure 2. A more formal definition is given now.

DEFINITION 3.2. Let $\Phi \in \mathcal{MT}$ such that $\#m(\Phi) > 1$ and let $l \in m(\Phi)$ be the node which corresponds to the final collision in the sense that $\bar{l} = \text{root}$ and $s_l = \tau(\Phi)$. The subtree $\Phi' = ((m', E'), (u', v'), (v'_k, s'_k, \nu'_k)_{k \in m'})$ is defined by

$$\begin{aligned}
 m' &= \{k \in m : l \leq_p k\}, \\
 E' &= \{\{k, k'\} : k, k' \in m', \{k, k'\} \in E\} \text{ with node data given by} \\
 (u', v') &= (u_l, v_l), (v'_k, \nu'_k, s'_k) = (v_k, \nu_k, s_k) \text{ if } l <_p k.
 \end{aligned}$$

The pruned tree $\bar{\Phi} = ((\bar{m}, \bar{E}), (u, v), (v_k, \nu_k, s_k)_{k \in \bar{m}})$ is defined by $\bar{m} = m \setminus m'$ and $\bar{E} = \{\{k, k'\} : k, k' \in \bar{m}, \{k, k'\} \in E\}$.

Recall that \mathcal{MT} is a metric space and denote for each $\Psi \in \mathcal{MT}$ by

$$B_h(\Psi) = \left\{ \Phi \in \mathcal{MT} : d(\Phi, \Psi) \leq \frac{h}{2} \right\},$$

the ball with diameter h centered at Ψ . For $0 < h < 2$ the ball $B_h(\Psi)$ is a $2d\#m(\Psi)$ dimensional smooth set.

DEFINITION 3.3. The standard Lebesgue measure on \mathcal{MT} is denoted by $d\lambda$.

We will now describe several probability measures on \mathcal{MT} to first describe the idealized distribution P_t , closely related to the Boltzmann equation, and then the empirical distributions \hat{P}_t , related to the annihilation flow. We collect several prop-

erties of these to prepare Proposition 5.5, which delivers the convergence of \hat{P}_t to P_t as $n \rightarrow \infty$.

4. Idealized distribution. The idealized distribution P_t is characterized by a differential equation (4.5). Before stating the equation we give a simple example which motivates the form and the analysis of the equation. Then we show that (4.5) admits a unique solution P_t . Finally we study the properties of P_t which will be instrumental when we demonstrate in section 5.1 that for each t the probability distribution P_t is very close to the empirical tree distribution \hat{P}_t , which is generated by the annihilation dynamics.

To motivate the analytical setting we consider first a simple example which illustrates the notation and the way semigroup theory applies. Recall that δ denotes the Dirac distribution and consider the linear system of differential equations

$$(4.1) \quad \begin{cases} \frac{d}{dt}u(t) = \mu u, & u(0) = 1, \\ \frac{\partial}{\partial t}v(s, t) = \delta(t - s)u + \mu v, & v(s, 0) = 0 \text{ for all } s, \end{cases}$$

with a parameter $\mu \leq 0$ and time-dependent variables $(u, v) \in X = \mathbb{R} \times M([0, \infty))$. The Banach space $M(\Omega)$ is the set of all finite unsigned measures, or alternatively, the dual space of $C(\Omega)$.

The solution is given by $u(t) = \exp(\mu t)$ and

$$v(s, t) = \begin{cases} \exp(\mu t) & \text{if } t \geq s \geq 0, \\ 0 & \text{else.} \end{cases}$$

The generator takes the form $L_t = (\delta_{(t-\cdot)} \mu \ 0)$; it is easy to see that L_t is stable for each $\mu < 0$, i.e., has a continuous resolvent for each $\lambda \geq 0$. Indeed, $(\lambda - \mu)f_1 = g_1$ and $-\delta(t - \cdot)f_1 + (\lambda - \mu)f_2 = g_2$ imply that

$$(4.2) \quad f_1 = \frac{1}{\lambda - \mu}g_1,$$

$$(4.3) \quad f_2 = \frac{1}{\lambda - \mu} \left(\frac{g_1}{\lambda - \mu} \delta(t - \cdot) + g_2 \right),$$

which is clearly a continuous map from X to X . In the case of the example, the operator L_t is actually continuous,

$$\begin{aligned} \|L_t f\|_X &= |\mu| \|f_1\| + \|f_1 \delta(t - \cdot) + \mu f_2\|_{M([0, \infty))} \\ &\leq |\mu| \|f_1\| + \underbrace{\|f_1\|}_{=1} \|\delta(t - \cdot)\|_{M([0, \infty))} + |\mu| \|f_2\|_{M([0, \infty))} \leq (1 + |\mu|) \|f\|_X. \end{aligned}$$

The operators below are not continuous and hence a more detailed analysis is required. A key result is that our approach delivers the existence of solutions $v \in L^1(\mathbb{R} \times [0, T])$. Indeed, assume for simplicity that $u, v \geq 0$, and the general case can be treated analogously. One obtains

$$(4.4) \quad \begin{aligned} \sup_{t \in [0, T]} \int_{\mathbb{R}} v(s, t) ds &= \sup_{t \in [0, T]} \int_0^t dr \int_{\mathbb{R}} ds \frac{\partial v}{\partial t}(s, r) \\ &\leq \sup_{t \in [0, T]} \int_0^t u(r) dr \leq \int_0^T |u(t)| dt < \infty, \end{aligned}$$

where the first inequality is obtained by estimating $\frac{\partial v}{\partial t}(s, t)$ from above by $\delta(s - t)u(t)$.

Now we consider a setting which is more closely linked with annihilation dynamics.

4.1. Existence and uniqueness of the idealized distribution. The idealized distribution P_t^a is characterized via a nonlinear Kolmogorov equation. The distribution P_0^a is supported on trees with only one node (the root). For $t > 0$ the support of the distribution P_t^a are those trees with the property $\tau < t$, i.e., the probability of all trees with $\tau > t$ is 0 by definition.

At each time t the gain term in the Kolmogorov equation is nonzero only on trees with the property $t = \tau$. This means that for a given tree Φ the time evolution $P_t^a(\Phi)$ is a nonnegative function which is zero for $t < \tau$ and nonincreasing for $t \geq \tau$. At time $t = \tau(\Phi)$ the function $t \mapsto P_t^a(\Phi)$ jumps instantaneously from 0 to a finite value which is determined by the probabilities of the subtrees $\bar{\Phi}$ and Φ' and the rate which depends only on $\bar{\Phi}$ and Φ' , but not on their probabilities. The loss term is nonzero only on trees with $t > \tau$; it is demonstrated in Lemma 4.1 that the loss term in (4.5) is the integral of the gain term.

We will show in section 5 that the time evolution of the probability distribution of trees in the empirical case satisfies a similar, albeit more involved, evolution equation. The similarity of the idealized and the empirical Kolmogorov equation is the prerequisite for the derivation of the analytical bounds in section 5.1 which deliver the closeness of the idealized and the empirical distribution of trees in the limit where a tends to 0.

The idealized tree distribution P_t^a satisfies a nonlinear Kolmogorov equation in the form

$$(4.5) \quad \begin{cases} \frac{\partial P_t^a}{\partial t} = \mathcal{Q}_t^a[P_t] = \mathcal{Q}_{t,+}^a[P_t^a] - \mathcal{Q}_{t,-}^a[P_t^a], \\ P_0^a(\Phi) = f_0(u(\Phi), v(\Phi)) \mathbf{1}_{\#m(\Phi)=1}, \end{cases}$$

where

$$(4.6) \quad \mathcal{Q}_{t,+}^a[P_t](\Phi) = \delta(t - \tau(\Phi)) P_t(\bar{\Phi}) P_t(\Phi') [(v - v') \cdot \nu]_+,$$

$$(4.7) \quad \mathcal{Q}_{t,-}^a[P_t](\Phi) = \mu_t^a[P_t](\Phi) P_t(\bar{\Phi})$$

with the convention $v = v(\Phi)$, $v' = v(\Phi')$, etc., and the loss rate $\mu_t^a[P] \in M(\mathcal{MT})$ is given by

$$(4.8) \quad \mu_t^a[P](\Phi) = \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} dP(\Psi) \delta(u - u(\Psi) + t(v - v(\Psi)) + a\nu) [(v - v(\Psi)) \cdot \nu]_+.$$

Formula (4.6) expresses the probability of Φ in terms of the subtrees $\bar{\Phi}$ and Φ' . The operator $\mathcal{Q}_{t,-}$ compensates the gain caused by $\mathcal{Q}_{t,+}$ with the result that \mathcal{Q}_t conserves the mass. From now on we will abbreviate the initial condition in (4.5) by using the convention

$$(4.9) \quad f_0(\Phi) = f_0(u(\Phi), v(\Phi)) \mathbf{1}_{\#m(\Phi)=1}.$$

Note that for a given tree Φ the operator L_t^a extracts the subtrees Φ' such that the roots of $\bar{\Phi}$ and Φ' collide at time t . The initial position $u(\Phi')$ varies with a as in (3.1) and provides the sole mechanism how L^a depends on a . The dependency on a will be mostly suppressed. The “idealized” distribution P_t is defined by $P_t = P_t^0$.

We will often use the quasi-linear form of the operator \mathcal{Q}_t in our analysis:

$$(4.10) \quad \mathcal{Q}_t[P, P', \mu](\Phi) = P(\bar{\Phi}) L_t[P', \mu](\Phi),$$

$$(4.11) \quad L_t[P', \mu](\Phi) = \delta(t - \tau(\Phi)) P'(\Phi') [(v - v') \cdot \nu]_+ - \mu_t(\Phi).$$

To see that \mathcal{Q}_t conserves the total probability we have to show that the delta distributions in (4.8) and (4.11) are equivalent.

LEMMA 4.1. *Let $\Phi \in \mathcal{MT}$. Then*

$$(4.12) \quad \delta(t - \tau(\Phi)) = \delta(u - u'(\Phi) + t(v - v'(\Phi)) + av(\Phi)),$$

i.e.,

$$\begin{aligned} & \int_{\mathcal{MT}} d\lambda(\Phi) g(\bar{\Phi}, \Phi') \delta(t - \tau(\Phi)) \\ &= \int_{\mathcal{MT}} d\lambda(\Phi) \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} d\lambda(\tilde{\Phi}) g(\Phi, \tilde{\Phi}) \delta(u(\Phi) - u(\tilde{\Phi}) + t(v(\Phi) - v(\tilde{\Phi})) + av) \end{aligned}$$

for all $g \in C_c(\mathcal{MT} \times \mathcal{MT})$, the set of continuous functions with compact support.

Proof. Let $\mathcal{MT}_* = \{\Phi \in \mathcal{MT} : u_{\text{root}} = 0\}$ and define for each $\Phi \in \mathcal{MT}_*$ and $u \in U$ the translated tree $\xi(\Phi, u) \in \mathcal{MT}$ as

$$\xi_l = \begin{cases} (u, v_{\text{root}}) & \text{if } l = \text{root,} \\ \Phi_l & \text{else.} \end{cases}$$

Then we find that the left-hand side of (4.12) can be rewritten as

$$\begin{aligned} & \int_{\mathcal{MT}} d\lambda(\Phi) g(\bar{\Phi}, \Phi') \delta(\tau - t) \\ &= \int_{\mathcal{MT}} d\lambda(\Phi) \int_{\mathcal{MT}_*} d\lambda(\tilde{\Phi}) \int_{S^{d-1}} d\nu \int_{[\bar{\tau}(\Phi), T]} d\tilde{\tau} \\ & \quad g\left(\Phi, \xi(\tilde{\Phi}, u + \tilde{\tau}(v - \tilde{v}) + av)\right) \delta(t - \tilde{\tau}) \\ &= \int_{\mathcal{MT}} d\lambda(\Phi) \int_{\mathcal{MT}_*} d\lambda(\tilde{\Phi}) \int_{S^{d-1}} d\nu g\left(\Phi, \xi(\tilde{\Phi}, u + t(v - \tilde{v}) + av)\right). \end{aligned}$$

Similarly we obtain for the right-hand side of (4.12)

$$\begin{aligned} & \int_{\mathcal{MT}} d\lambda(\Phi) \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} d\lambda(\tilde{\Phi}) g(\Phi, \tilde{\Phi}) \delta(u(\Phi) - u(\tilde{\Phi}) + t(v(\Phi) - v(\tilde{\Phi})) + av) \\ &= \int_{\mathcal{MT}} d\lambda(\Phi) \int_{S^{d-1}} d\nu \int_{\mathcal{MT}_*} d\lambda(\tilde{\Phi}) g\left(\Phi, \xi(\tilde{\Phi}, u + t(v - \tilde{v}) + av)\right). \end{aligned}$$

Hence, both sides of (4.12) coincide and the proof is finished. \square

An immediate consequence of Lemma 4.1 is that the average of $\mathcal{Q}_t[P, P']$ is zero in the sense that

$$(4.13) \quad \int_{\mathcal{MT}} d\mathcal{Q}_t[P, P'](\Phi) g(\bar{\Phi}) = 0 \quad \text{for all } g \in C_c(\mathcal{MT}).$$

The relation to the Boltzmann equation will become apparent in Proposition 4.9. To study the existence of solutions of (4.5), we first introduce the appropriate function spaces. We define spaces of integrable functions on \mathcal{MT} with general weights

$$(4.14) \quad X_\ell := \{f \in M(\mathcal{MT}) \mid \|f\|_{X_\ell} < \infty \text{ and } f(\cdot \mid \tau = t) \in L^1(\mathcal{MT}_t) \quad \forall t \geq 0\}$$

with

$$\|f\|_{X_\ell} = \sup \left\{ \int_0^T ds \int_{\mathcal{MT}} d|f|(\Phi) w(u + sv, s) (1 + |v|)^\ell \mid \int_{U \times [0, T]} du ds w(u, s) = 1 \right\},$$

and let $X = X_1$. Note that X_ℓ is a Banach space but it is not a subset of $L^1(\mathcal{MT})$ because the τ -marginal can have concentrations.

Remark 4.2. If $P \in X_\ell$, $P|_{\{\Phi:\#m(\Phi)=1\}} \in L^1(U \times \mathbb{R}^d)$ and the τ -marginal is in $L^1((0, T])$, then $P \in L^1(\mathcal{MT})$.

To see that X_ℓ is a Banach space we suppose that f_m is a Cauchy sequence in X_ℓ ; then $f_m \rightarrow f$ in $M(\mathcal{MT})$. Since the sequence f_m converges it is also tight. The absolute continuity follows from the disintegration theorem [DM78], which provides the existence of $\sigma \in M([0, T])$ and a family of measures $f_t = f(\cdot | \tau = t) \in L^1(\mathcal{MT}_t)$ such that for all $g \in C(\mathcal{MT})$ the formula

$$\int_{\mathcal{MT}} g \, df = \int_0^T d\sigma(t) \int_{\mathcal{MT}_t} df_t(\Phi) g(\Phi)$$

holds. Then for $E_t \subset \mathcal{MT}_t$ of measure zero $f_t(E_t) = 0$ for σ -almost every t , such that we have $f \in X_\ell$ after a modification on a set of measure zero.

After this preparation we can derive an existence and uniqueness result for the linearized evolution, where we fix the second argument in (4.10). The distribution P and the operator \mathcal{Q} depend on a , the dependency is not shown for the sake of notational convenience.

LEMMA 4.3. *For each $P' \in C^0([0, T], X)$ the operator $\mathcal{Q}_t[\cdot, P'_t]$ is the generator of a strongly continuous evolution $U(s, t)$ on X , i.e., there exists a unique solution of the equation*

$$(4.15) \quad \frac{\partial}{\partial t} P_t = \mathcal{Q}_t[P_t, P'_t], \quad P_0 = f_0.$$

For each $t > 0$ the solution P_t has the following properties:

- (i) P_t has a density, i.e., $P_t \in L^1(\mathcal{MT})$.
- (ii) P_t is nonnegative, i.e., $\int_\Omega dP_t(\Phi) \geq 0$ for all $\Omega \subset \mathcal{MT}$ measurable.
- (iii) P_t is normalized, i.e., $\int_{\mathcal{MT}} dP_t(\Phi) = 1$.
- (iv) The Lagrangian root marginal $\pi[P_t]$ which is defined by

$$\int_{U \times \mathbb{R}^d} d\pi[P](u, v) g(u, v) = \int_{\mathcal{MT}} dP(\Phi) g(u(\Phi), v(\Phi)) \text{ for all } g \in C_c(U \times \mathbb{R}^d)$$

is independent of a and t , i.e., $\pi[P_t] = f_0$.

Remark 4.4. As a consequence of Lemma 4.3(iv) we obtain the following formula for the collision rate which involves only f_0 but not P_t :

$$(4.16) \quad \mu_t[P_t](u, v) = \mu_t(u, v) = \int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} d\nu' f_0(u + t(v - v') + a\nu, v') [(v - v') \cdot \nu]_+.$$

Proof of Lemma 4.3. We show that $\mathcal{Q}_t[\cdot, P']$ generates an evolution on X with X_2 being a subset of the domain of the unbounded operator $\mathcal{Q}_t[\cdot, P'_t]$; for this we use general results of [Paz83, Chapter 5]. The aim is to prove the existence of an evolution operator, which is the nonautonomous version of a semigroup. We study the resolvent equation

$$(4.17) \quad R_\lambda h = g$$

with $R_\lambda h = \lambda h - \mathcal{Q}_t[h, P']$ for $\lambda > 0$ and $g \in X$. It suffices to establish the existence of nonnegative solutions h if $g \geq 0$. Indeed, for general h we can decompose h and

g into positive and negative parts: $h = h_+ - h_-$, $g = g_+ - g_-$; if $R_\lambda h_\pm = g_\pm$, then $R_\lambda(h_+ - h_-) = g_+ - g_-$. We consider two separate cases depending on whether the time coincides with the final collision time of a tree.

Then for $t \neq \tau(\Phi)$, we obtain

$$(4.18) \quad \lambda h(\Phi) + \mu(\Phi)h(\Phi) = g(\Phi),$$

i.e.,

$$(4.19) \quad h(\Phi) = \frac{1}{\lambda + \mu(\Phi)}g(\Phi).$$

For $t = \tau(\Phi)$ we seek a solution to

$$(4.20) \quad (\lambda h - \mathcal{Q}_t[h, P'])(\Phi) = \lambda h(\Phi) + \mu(\Phi)h(\Phi) - h(\bar{\Phi})P'(\Phi')((v - v_{\text{root}}(\Phi')) \cdot \nu)_+ = g(\Phi).$$

We are using that $t \neq \tau(\bar{\Phi})$ for the pruned tree $\bar{\Phi}$, such that we can use (4.19) for the $h(\bar{\Phi})$ expression. Then the solution to (4.20) is given by

$$(4.21) \quad h(\Phi) = \frac{1}{\lambda + \mu(\Phi)} \left(g(\Phi) + g(\bar{\Phi}) \frac{P'(\Phi')(v - v') \cdot \nu_+}{\lambda + \mu_t(\bar{\Phi})} \right),$$

where $v' = v_{\text{root}}(\Phi')$. The key observation in (4.19) and (4.21) is that h is nonnegative for nonnegative g and $\lambda > 0$. Hence for nonnegative h we have

$$\|h\|_X = \sup \left\{ \int_{\mathcal{MT}} dh(\Phi) (1 + |v|) w(u + sv, s) \mid \int_{U \times [0, T]} du ds w(u, s) = 1 \right\}$$

(without the $|\cdot|$ bars). Then we write (4.17) as $h = \frac{1}{\lambda}(g + \mathcal{Q}_t[h, P'])$ and we obtain that for each $w \in L^1(U \times [0, T])$ the equation

$$\begin{aligned} & \int_0^T ds \int_{\mathcal{MT}} dh(\Phi) (1 + |v|) w(u + sv, s) \\ &= \frac{1}{\lambda} \int_0^T ds \int_{\mathcal{MT}} dg(\Phi) (1 + |v|) w(u + sv, s) \\ & \quad + \frac{1}{\lambda} \int_0^T ds \int_{\mathcal{MT}} d\mathcal{Q}_t[h, P'](1 + |v|) w(u + sv, s) \\ &= \frac{1}{\lambda} \int_0^T ds \int_{\mathcal{MT}} dg(\Phi) (1 + |v|) w(u + sv, s) \end{aligned}$$

holds. The first equation is due to (4.17), and the second equation follows from (4.13). Nonpositive right-hand-sides h can be treated analogously. Thus

$$\|h\|_X = \frac{1}{\lambda} \|g\|_X.$$

This shows that $\mathcal{Q}_t[\cdot, P'_t]$ is a stable family of generators with exponential growth rate 0 and bound $M = 1$. Furthermore as (4.13) also holds when restricting to Φ with given root data, then also $h \in X_\ell$ if $g \in X_\ell$.

We will demonstrate now that for each $h \in X_2$ (cf. (4.14)), $t \in [0, T]$, and $P' \in X$,

$$(4.22) \quad \|\mathcal{Q}_t[h, P']\| \leq 2\kappa_d \|h\| \mu_t[P'].$$

Indeed, for fixed $t \in [0, T]$ one obtains that

$$\begin{aligned} & \| \mathcal{Q}_t[h, P'] \|_X \\ & \leq \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|) w(u + sv, s) \\ & \quad \times \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} dP'(\Phi') \delta(u - u' + t(v - v') + a\nu) (v - v') \cdot \nu_+ \\ & \quad + \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) \mu_t[P'](\Phi) (1 + |v|) w(u + sv, s). \end{aligned}$$

It is immediate that the first term and the second term coincide:

$$\begin{aligned} & \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|) w(u + sv, s) \\ & \quad \times \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} dP'(\Phi') \delta(u - u' + t(v - v') + a\nu) (v - v') \cdot \nu_+ \\ & = \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|) w(u + sv, s) \mu_t[P'](\Phi), \end{aligned}$$

and thus inequality (4.22) is established. Hence, we obtain for each $t \in [0, T]$

$$\begin{aligned} & 2 \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|) w(u + sv, s) \int_{\mathcal{MT}} d|L_t[P'](\Phi')| \\ & \leq 2 \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|) w(u + sv, s) \\ & \quad \times \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} dP'_t(\Phi'') \delta(u - u'' + t(v - v'') + a\nu) (v - v'') \cdot \nu_+ \\ & \leq 2 \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|) w(u + sv, s) \\ & \quad \times \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} dP'_t(\Phi'') \delta(u - u'' + t(v - v'') + a\nu) (1 + |v|)(1 + |v''|) \\ & \leq 2 \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|)^2 w(u + sv, s) \\ & \quad \times \sup \left\{ \int_{S^{d-1}} d\nu \int_{\mathcal{MT}} dP'_t(\Phi'') \tilde{w}(u - u'' + t(v - v'') + a\nu) (1 + |v''|) : \right. \\ & \quad \left. \int_U \tilde{w}(u) du = 1 \right\} \\ & \leq 2\kappa_d \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|)^2 w(u + sv, s) \\ & \quad \times \sup \left\{ \int_0^T ds \int_{\mathcal{MT}} P'_t(\Phi'') (1 + |v''|) \tilde{w}(u'' + sv'', s) : \right. \\ & \quad \left. \int_{U \times [0, T]} du ds \tilde{w}(u, s) = 1 \right\} \\ & = 2\kappa_d \int_0^T ds \int_{\mathcal{MT}} d|h|(\bar{\Phi}) (1 + |v|)^2 w(u + sv, s) \|P'\|_X. \end{aligned}$$

This implies that

$$(4.23) \quad \|\mathcal{Q}_t[h, P']\|_X \leq 2\kappa_d \|h\|_{X_2} \|P'\|_X,$$

hence $X_2 \subset D(\mathcal{Q}_t[\cdot, P'_t])$.

We are now in a position to check conditions $(H_1), (H_2), (H_3)$ in [Paz83, Theorem 5.3.1]. The first two conditions are satisfied as $\mathcal{Q}_t[\cdot, P'_t]$ is a stable family of generators with exponential growth rate 0 and bound 1 both on X and $Y = X_2$. By (4.23) and linearity in the second argument we also obtain that $t \mapsto \mathcal{Q}_t[\cdot, P'_t]$ is continuous in the $\|\cdot\|_{Y \rightarrow X}$ norm as long as $t \mapsto P'_t$ is continuous in X . Then there exists a unique evolution system $U(t, t_0)$ by [Paz83, Theorem 5.3.1] on X .

Now we demonstrate that $P_t \geq 0$. The construction of $U(t, s)$ is based on repeated applications of $(\text{Id} - (\Delta t)\mathcal{Q}_s[\cdot, P'_s])^{-1}$ with $\Delta t > 0$ and $s \in [0, T]$; all these operators are multiples of the resolvent in (4.17). Hence they map positive functions to positive functions by the observation after (4.21). Thus we have $P_t \geq 0$.

Together with (4.13) this implies that for each $t \geq 0$ the measure P_t characterizes a probability distribution on \mathcal{MT} .

Next we show that the measure P_t has a density with respect to the Lebesgue measure for all $t \geq 0$ by adapting (4.4). We note that P_t has a density on the trees just consisting of the roots due to absolute continuity of f_0 . Following Remark 4.2 it suffices then to show that there exists a function $h \in L^1([0, T])$ with the property

$$(4.24) \quad \int_{\{\Phi \in \mathcal{MT} : \#m(\Phi) > 1\}} dP_t(\Phi) g(\tau(\Phi)) = \int_0^t dh(s) g(s) \quad \text{for all } g \geq 0, g \in L^\infty((0, T]).$$

First note that (4.24) is a consequence of the stronger bound

$$\int_{\{\Phi \in \mathcal{MT} : \#m(\Phi) > 1\}} dP_t(\Phi) g(\tau(\Phi)) \leq C \|g\|_{L^1((0, T])} \quad \text{for all } g \geq 0, g \in L^1((0, T]),$$

which gives the existence of $h \in L^\infty([0, T]) \subset L^1([0, T])$ in (4.24). As P_t solves (4.15) strongly, $\partial_t P \in C^0((0, T), X)$ and we have that for each $t \in [0, T]$ by $P_t = f_0 + \int_0^t \partial_s P_s ds$

$$\begin{aligned} \int_{\{\Phi \in \mathcal{MT} : \#m(\Phi) > 1\}} dP_t(\Phi) g(\tau(\Phi)) &\leq \int_0^t ds \int_{\mathcal{MT}_s} d\partial_s P_s(\Phi) g(s) \\ &\leq \|g\|_{L^1([0, T])} \sup_{t \in [0, T]} \|\partial_t P_t\|_X, \end{aligned}$$

where we used $\partial_s P_s(\Phi)$ is only positive for $s = \tau(\Phi)$.

Finally by letting $\varphi(\bar{\Phi}) = g(u(\bar{\Phi}), v(\bar{\Phi})) = g(u(\Phi), v(\Phi))$, part (iv) follows from (4.13). \square

We will also need an L^1 version of (4.23).

Remark 4.5. Lemma 4.3 also holds if X is replaced with the Banach space

$$Z = \{f \in L^1(\mathcal{MT}) \mid \|f\|_Z < \infty\}$$

with $\|f\|_Z = \int_{\mathcal{MT}} d|f|(\Phi) (1 + |v|)$ since estimate (4.23) follows with X replaced by Z , as $\mu_t(u, v)$ is an $L^1(\mathbb{T}^d)$ function for the argument $(u + tv)$, whereas $f \in X_2$ is $L^\infty(\mathbb{T}^d)$ for the same argument, such that the estimate of the product term $\mathcal{Q}_t[f, P']$ follows by the Hölder inequality. Note, however, that Z is not a suitable space for establishing the existence of a nonlinear semigroup.

Using a hierarchy of approximations we are now able to obtain the idealized distributions.

PROPOSITION 4.6.

- (i) For each $a > 0$ the nonlinear Kolmogorov equation (4.5) has a unique solution $P_t^a \in C^1([0, T], X) \cap C^0([0, T], X_2)$ for every f_0 satisfying (1.5).
- (ii) For given initial data and for each t the measure P_t^a converges to $P_t = P_t^0$ in Z as $a \rightarrow 0$.

Proof. As the role of a is not relevant for (i) we will not show the dependency on a in this part of the proof. We prove (i) by approximating P_t by a sequence of probability measures $P_{t,k}$ which are defined recursively by the equation

$$(4.25) \quad P_{t,1} = f_0,$$

$$(4.26) \quad \frac{\partial P_{t,k}}{\partial t} = \mathcal{Q}_t[P_{t,k}, P_{t,k-1}], \quad P_{0,k} = f_0,$$

where the convention (4.9) has been used. The existence of an evolution operator for (4.25), (4.26) if $P_{t,k} \in C^1([0, T], X) \cap C^0([0, T], Y)$ is a consequence of Lemma 4.3. To have classical solutions of the operator equation we have to use some more semigroup theory. The evolution system in Lemma 4.3 is constructed through an implicit Euler approximation, i.e., using a resolvent as in (4.17). As the resolvents leave Y invariant, $U(t, t_0)$ maps Y and also any other X_ℓ to itself, giving condition (E_4) in [Paz83, Theorem 5.4.3].

To check the strong continuity condition in Y , condition (E_5) , we start with initial data in $f \in X_2$ and use the previous results with X replaced by X_2 and Y by X_3 . Then Theorem 5.3.1 in [Paz83] implies that there is a unique Y -valued solution of

$$(4.27) \quad \begin{cases} \frac{\partial P_t}{\partial t} = \mathcal{Q}_t[P_t, P_t'], \\ P_0 = f_0. \end{cases}$$

Replacing P' by $P_{t,k-1}$ and P_t by $P_{t,k}$ gives that $P_{t,k} \in C^1([0, T], X) \cap C^0([0, T], Y)$ for all $k \in \mathbb{N}$ by induction.

Next we will prove that $P_t = \lim_{k \rightarrow \infty} P_{t,k}$ exists by showing that the solutions of the nonautonomous linear equation (4.27) are a contraction of $P' \in C^0([0, T], X)$ with respect to the norm

$$\|P'\|_\rho = \sup_{t \in [0, T]} \exp(-\rho t) \|P_t'\|_X.$$

To consider a solution of (4.27), we replace P' with \tilde{P}'_t :

$$(4.28) \quad \begin{cases} \frac{\partial \tilde{P}_t}{\partial t} = \mathcal{Q}_t[\tilde{P}_t, \tilde{P}'_t], \\ \tilde{P}_0 = f_0; \end{cases}$$

then $P_t - \tilde{P}_t$ satisfies

$$(4.29) \quad \begin{cases} \frac{\partial P_t}{\partial t} - \frac{\partial \tilde{P}_t}{\partial t} = \mathcal{Q}_t[P_t - \tilde{P}_t, \tilde{P}'_t] + \mathcal{Q}_t[P_t, P_t' - \tilde{P}'_t], \\ P_0 - \tilde{P}_0 = 0. \end{cases}$$

Then the strong solution $P_t - \tilde{P}_t \in C^1([0, T], X) \cap C^0([0, T], X_2)$ constructed above can be represented as a mild solution of (4.29). Lemma 4.3 gives the evolution $U(t, s)$ generated by \tilde{P}' ; thus

$$P_t - \tilde{P}_t = \int_0^t U(t, s) \mathcal{Q}_s[P_s, P_s' - \tilde{P}'_s] ds.$$

Then (4.23), the boundedness of $U(t, s)$, and the fact that $\|P_s\|_{X_2} \leq K_\infty$ for all $s \geq 0$ give the estimate

$$\begin{aligned} \|P - \tilde{P}\|_\rho &= \sup_{t \in [0, T]} \exp(-\rho t) \left\| \int_0^t U(t, s) \mathcal{Q}_s[P_s, P'_s - \tilde{P}'_s] ds \right\|_X \\ &\leq 2\kappa_d \sup_{t \in [0, T]} \exp(-\rho t) \int_0^t \|P_s\|_{X_2} \|P'_s - \tilde{P}'_s\|_X ds \\ &\leq 2\kappa_d \sup_{t \in [0, T]} \exp(-\rho t) \int_0^t K_\infty \exp(\rho s) \|P' - \tilde{P}'\|_\rho ds \\ &\leq \frac{2\kappa_d K_\infty}{\rho} \left(\sup_{t \in [0, T]} (1 - \exp(-\rho t)) \right) \|P' - \tilde{P}'\|_\rho, \end{aligned}$$

i.e., this is a contraction for $\rho > 2\kappa_d K_\infty$. Thus $P_{\cdot, k}$ converges in $C^0([0, T], X)$ to a unique fixed point P . Setting $P'_t = P_t$ in (4.27) and using Lemma 4.3 then gives the desired regularity.

To prove (ii), we reintroduce the parameter a with the convention that $P_t = P_t^0$ and $\mathcal{Q}_t = \mathcal{Q}_t^0$. The difference $P_t^a - P_t$ satisfies

$$\frac{\partial P_t^a}{\partial t} - \frac{\partial P_t}{\partial t} = \mathcal{Q}_t[P_t^a - P_t, P_t] + \mathcal{Q}_t^a[P_t^a, P_t^a - P_t] + P_t^a(L_t^a[P_t] - L_t[P_t]).$$

Denoting the evolution generated by $\mathcal{Q}_t[\cdot, P_t]$ as $U(t, s)$, we obtain

$$P_t^a - P_t = \int_0^t U(t, s) \{ \mathcal{Q}_s^a[P_s^a, P_s^a - P_s] + P_s^a(L_s^a[P_s] - L_s[P_s]) \} ds.$$

By Remark 4.5 we obtain the bound

$$(4.30) \quad \|\mathcal{Q}_t^a[P, P']\|_Z \leq 2\kappa_d \|P\|_{X_2} \|P'\|_Z.$$

Using (4.30) and that $U(t, s)$ is a bounded operator on Z we arrive at

$$\begin{aligned} \|P_t^a - P_t\|_Z &= \int_0^t \|U(t, s) \mathcal{Q}_s^a[P_s^a, P_s^a - P_s]\|_Z ds + \int_0^t \|P_s^a(L_s^a[P_s] - L_s[P_s])\|_Z ds \\ &\leq C \int_0^t \{ \|P_s^a\|_{X_2} \|P_s^a - P_s\|_Z + \|P_s^a\|_{X_2} \|L_s^a[P_s] - L_s[P_s]\|_Z \} ds. \end{aligned}$$

Due to strong continuity of spatial shifts in the L^1 norm used for Z , the last term converges to 0 as $a \rightarrow 0$. Gronwall's inequality gives the required convergence in (ii). \square

Remark 4.7. The existence result Proposition 4.6(i) delivers a tightness bound on the number of nodes $\#m(\Phi)$ of trees Φ in P_t of the following form: there exists a function $M(\varepsilon)$ such that

$$(4.31) \quad P_t(\{\#m(\Phi) \geq M(\varepsilon)\}) \leq \varepsilon$$

for all $\varepsilon > 0$.

4.2. Properties of the idealized distribution. It is interesting to note that P_t is a historical Markovian process (cf. [DP91]) in the following sense:

$$(4.32) \quad \int_{\mathcal{MT}} dP_t(\Phi) g(\Phi) = \int_{\mathcal{MT}} dP_s(\Phi) g(\Phi)$$

for all $s \leq t$ and $g \in C(\Phi)$ with the property $g(\Phi) = g(\text{pr}_s(\Phi))$ for all $\Phi \in \mathcal{MT}$, where

$$\begin{aligned} \text{pr}_t(\Phi) &= ((m_t, E), (u, v), (s_l, \nu_l, v_l)_{l \in m_t}), \\ m_t &= \{l \in m : \exists l' \geq_p l \text{ such that } \bar{l}' = \text{root and } s_{l'} \leq t\} \cup \{\text{root}\} \end{aligned}$$

denotes the stripped tree where all collisions after time t are removed. Equation (4.32) shows that no information is lost. As this fact is irrelevant for our purposes we will not give a proof.

On the other hand, it is possible to find Markovian random variables with constant complexity. As an intermediate step toward constructing Markovian random variables we show that the subtrees with collision times $t \geq s$ are Markovian with respect to $(u(s), v(s))$.

DEFINITION 4.8. *For each tree $\Phi \in \mathcal{MT}$ the random variable $\beta(\Phi) \in \{0, 1\}$ is defined recursively by*

$$(4.33) \quad \beta(\Phi) = \begin{cases} 1 & \text{if } \#m(\Phi) = 1, \\ \beta(\bar{\Phi})(1 - \beta(\Phi')) & \text{else.} \end{cases}$$

The random variable $\beta(\Phi)$ is the indicator function of those trees where the root particle has not undergone a collision. We will show now that the expectation of this observable satisfies a closed evolution equations.

PROPOSITION 4.9. *Consider $P_t = P_t^0$ as in Proposition 4.6. The marginal distribution*

$$f_t(u, v) = P_t(\beta = 1 \text{ and } (u_{\text{root}}, v_{\text{root}}) = (u - tv, v))$$

satisfies the closed equation

$$(4.34) \quad \frac{\partial f}{\partial t} + v \cdot \nabla_u f = -Q_-[f, f], \quad f_{t=0}(u, v) = f_0(u, v)$$

in the weak sense, where $Q_-[\cdot, \cdot]$ is defined in (2.6).

The appearance of the transport term $v \cdot \nabla_u f$ is a result of the change from Lagrangian to Eulerian coordinates: P_t provides the distribution of the initial positions and velocities and $f_t(u, v)$ characterize the densities of particles with velocity v at position u at time t . An analogous statement is also true if $a > 0$, but this is irrelevant for our purposes.

Proof. Let $g \in C^1(U \times \mathbb{R}^d)$ be a test function. Then

$$\begin{aligned} & \frac{d}{dt} \int df_t(u, v) g(u, v) \\ &= \frac{d}{dt} \int dP_t(\Phi) \beta(\Phi) g(u + tv, v) \\ &= \int_{\mathcal{MT}} dQ_t[P_t, P_t](\Phi) \beta(\Phi) g(u + tv, v) + \int dP_t(\Phi) \beta(\Phi) v \cdot \nabla g(u + tv, v) \\ &= \int_{\mathcal{MT}} dQ_t[P_t, P_t](\Phi) \beta(\Phi) g(u + tv, v) + \int_{U \times \mathbb{R}^d} df_t(u + tv, v) v \cdot \nabla g(u + tv, v) \\ &= I_1 + I_2. \end{aligned}$$

A change of variables in I_2 yields that

$$I_2 = \int_{U \times \mathbb{R}^d} df_t(u, v) v \cdot \nabla g(u, v).$$

We analyze now I_1 . The definition of β implies that

$$\begin{aligned} I_1 &= \int_{\mathcal{MT}} d\mathcal{Q}_t[P_t, P_t](\Phi) \beta(\bar{\Phi}) g(u + tv, v) \mathbf{1}_{\tau < t} \\ &\quad + \int_{\mathcal{MT}_t} d\mathcal{Q}_t[P_t, P_t](\Phi) \beta(\Phi) g(u + tv, v) \beta(\bar{\Phi}) (1 - \beta(\Phi')) \\ &= \int_{\mathcal{MT}} d\mathcal{Q}_t[P_t, P_t](\Phi) \beta(\bar{\Phi}) g(u + tv, v) \\ &\quad - \int_{\mathcal{MT}_t} d\mathcal{Q}_t[P_t, P_t](\Phi) g(u + tv, v) \beta(\bar{\Phi}) \beta(\Phi'). \end{aligned}$$

Thanks to formula (4.13) the first term vanishes. The definition of \mathcal{Q}_t implies that

$$\begin{aligned} &\int_{\mathcal{MT}_t} d\mathcal{Q}_t[P_t, P_t](\Phi) \beta(\bar{\Phi}) \beta(\Phi') g(u + tv, v) \\ &= \int_{U \times \mathbb{R}^d} dQ_-[f_t, f_t](u + tv, v) g(u + tv, v) = \int_{U \times \mathbb{R}^d} dQ_-[f_t, f_t](u, v) g(u, v). \end{aligned}$$

The last equality follows from a change of variables. Putting everything back together we find that

$$\begin{aligned} &\frac{d}{dt} \int df_t(u, v) g(u, v) \\ &= - \int_{U \times \mathbb{R}^d} dg(u, v) Q_-[f_t, f_t](u, v) + \int_{U \times \mathbb{R}^d} df_t(u, v) v \cdot \nabla_u g(u, v) \end{aligned}$$

for all test functions $g \in C^1(U \times \mathbb{R}^d)$, which is the claim. \square

The link between the mild solutions in Proposition 4.6 and weak solutions in Proposition 4.9 is provided by the following proposition. We derive a formula that can be evaluated for a wide class of measures.

PROPOSITION 4.10. *Let $f_t \in L^2(U, L^1(\mathbb{R}^d))$.*

(i) *The equation*

$$f_t(u, v) = f_0(u - tv, v) \exp \left(- \int_0^t \int_{\mathbb{R}^d} f_s(u - (t-s)v, v') \kappa_d |v - v'| dv' ds \right), \tag{4.35}$$

is satisfied for all $t \in (0, T)$ if and only if f_t is the unique mild solution of (2.5).

(ii) *Equation (4.35) implies that f_t is also a distributional solution. Furthermore, every distributional solution with $Q_-[f, f] \in L^1((0, T) \times U, L^1_{1+|v|}(\mathbb{R}^d))$ is a mild solution.*

Proof. An equivalent formulation for f_t being a mild solution (2.7) is

$$(4.36) \quad f^\#(u, v, t) - f_0(u, v) = - \int_0^t (Q_-[f_s, f_s])^\# ds$$

with $h_t^\#(u, v) = h_t(u + tv, v)$.

First we show (i). Let f_t be a mild solution; then

$$\begin{aligned}
 f_t^\#(u, v) - f_0(u, v) &= - \int_0^t (L[f_s]f_s)^\#(u, v) ds = - \int_0^t (L[f_s])^\#(u, v) f_s^\#(u, v) ds \\
 (4.37) \qquad \qquad \qquad &= - \int_0^t g(u, v, s) f_s^\# ds
 \end{aligned}$$

with

$$g(u, v, s) = (L[f_s])^\#(u, v) = \int_{\mathbb{R}^d} f_s(u + sv, v') \kappa_d |v - v'| dv'$$

such that $g(u, \cdot, s) \frac{1}{1+|\cdot|} \in L^2_{loc}(\mathbb{R}^d)$. A solution to (4.37) is given by

$$(4.38) \qquad \qquad \qquad f_t^\#(u, v) = \exp\left(- \int_0^t g(u, v, s) ds\right) f_0(u, v),$$

as for each u, v the equation decouples to a single scalar ordinary differential equation, so f_t fulfills (4.35). For fixed u , the integral equation (4.38) has a unique solution in $C^0([0, T], L^1_{1+|\cdot|}(\mathbb{R}^d))$ by a simple contraction argument for finite times as in the spatially homogeneous case [MT10, Lemma 5]. This observation also shows that f_t given by (4.35) is a mild solution, completing the proof of (i).

For part (ii), we observe that mild and distributional solutions coincide following [Ba77] for

$$\partial_t f + v \cdot \nabla_u f = h$$

as long as $h = -Q_-[f_t, f_t] \in L^1((0, T), L^1_{1+|\cdot|}(U \times \mathbb{R}^d))$, which immediately shows the second part of (ii). For solutions given by (4.35), we first observe $0 \leq f_t(u, v) \leq f_0(u - tv, v)$ by (4.38). Hence

$$0 \leq Q_-[f_t, f_t] = L[f_t](u, v) f_t(u, v) \leq L[f_0(u - tv, \cdot)](v) f_0(u - tv, v),$$

such that $Q_-[f_t, f_t] \in L^1_{1+|\cdot|}(U \times \mathbb{R}^d)$ as

$$\|L[f_0 \circ \varphi] f_0 \circ \varphi\|_{L^1_{1+|\cdot|}} = \|L[f_0] f_0\|_{L^1_{1+|\cdot|}} < \infty$$

with $\varphi(u, v) = (u - tv, v)$. The last equation is due to (1.5). This completes the proof. \square

Remark 4.11. Interestingly there exists an explicit solution of the nonlinear Kolmogorov equation (4.5), but this fact is not relevant for our analysis. The corresponding expressions were used in the analysis of the idealized distribution in [MT10]. They can be obtained from (4.16) and explicit calculation of the rates.

Let $\Omega \subset \mathcal{MT}$ and $t \in [0, \infty)$. Then the idealized distribution is given by

$$(4.39) \qquad \qquad \qquad P_t(\Omega) = \int_{\Omega} \exp\left(-I_t(\Phi)\right) d\lambda(\Phi),$$

where the integrated collision rate is I_t is defined recursively,

$$\begin{aligned}
 I_t(\Phi) &= \int_0^t \Gamma_s(\Phi) ds + \sum_{(s, \Psi) \in \{(\tau(\Phi), \Phi'), (\tau(\bar{\Phi}), (\bar{\Phi})'), \dots\}} I_s(\Psi), \\
 \Gamma_t(\Phi) &= \int_{\mathbb{R}^d} dv' \int_{S^{d-1}} d\nu f_0(u - t(v - v') - a\nu, v') [(v - v') \cdot \nu]_+, \\
 d\lambda(\Phi) &= f_0(u, v) du dv \prod_{l \in m(\Phi) \setminus \{\text{root}\}} (f_0(u_l, v_l) [(v_l - v_l) \cdot \nu_l]_+ ds_l d\nu_l dv_l),
 \end{aligned}$$

where the initial positions u_l are defined by formula (3.1).

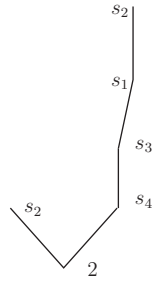


FIG. 3. A tree with recollision is obtained for the particle with label 2 in Figure 2. The particle colliding with particle 2 at time s_2 appears twice in the tree.

We end the section by introducing trees without “recollisions.” Despite particles undergoing at most one collision, an effect akin to recollision occurs when the same particle appears in multiple positions within the same collision tree (see, e.g., Figure 3). A particle will appear again when it has an intersection with a particle that is not its parent or one of its children. We introduce a set of good trees, which have uniform bounds on the maximal velocities on the number of nodes in the tree and which do not have recollisions.

DEFINITION 4.12. A tree Φ is recollision free on the time interval $[0, t]$ at diameter a if

$$|u_l^a + sv_l - u_{l'}^a - sv_{l'}| > a \text{ for all } 0 < s < t \text{ and } l, l' \in m(\Phi) \text{ such that } \{l, l'\} \notin E(\Phi). \tag{4.40}$$

For any pair of monotonic functions $M(a), V(a)$ such that $\lim_{a \rightarrow 0} M(a) = \lim_{a \rightarrow 0} V(a) = +\infty$ the set of good trees is defined as

$$\mathcal{G}(a_0) = \left\{ \Phi \in \mathcal{MT} : \#m(\Phi) \leq M(a_0) \text{ and } \max_{l \in m(\Phi)} |v_l| < V(a_0) \text{ and} \right. \\ \left. \min_{l \in m(\Phi) \setminus \text{root}} \nu_l \cdot (v_l - v_{l'}) > 0 \text{ and (4.40) holds for all } t = s_i \text{ and } a \in [0, a_0] \right\}. \tag{4.41}$$

Note that thanks to the monotonicity of V and M the set $\mathcal{G}(a_0)$ is monotonic in a_0 .

LEMMA 4.13. The good trees have almost full measure, i.e.,

$$\lim_{a \rightarrow 0} P_t^a(\mathcal{G}(a)) = \lim_{a \rightarrow 0} P_t(\mathcal{G}(a)) = 1. \tag{4.42}$$

Proof. We first show that $\mathcal{G}(0)$ is a set of measure 1. The only nontrivial condition is (4.40) with $a = 0$. Let $(\sigma, E) \in \mathcal{T}$ be a tree and define $\mathcal{MT}(\sigma, E) := \{\Phi \in \mathcal{MT} : m(\Phi) = \sigma, E(\Phi) = E\}$. Recall (3.1), which provides for each $l \in \sigma$ a recursive formula for the initial position u_l . We will write $u_l(s_l, v_l)$ to emphasize the dependency of the initial position on the collision time s_l and velocity v_l .

The dimension of $\mathcal{MT}(\sigma, E)$ is $(2d)^{\#\sigma}$ as the nodes are parameterized by $(u, v) \in \mathbb{T}^d \times \mathbb{R}^d$ for the root and by $(s, \nu, v) \in \mathbb{R}^d \times S^{d-1} \times [0, T]$. On the other hand, for $a = 0$ any pair $l, l' \in \sigma$ and fixed $(u_l, v_l) \in U \times \mathbb{R}^d$ and fixed $(u_{l'}, v_{l'}) \in U \times \mathbb{R}^d$, the subset of $\mathcal{MT}(\sigma, E)$ with

$$\{(v_{l'}, \nu_{l'}, s_{l'}) \in \mathbb{R}^d \times S^{d-1} \times [0, T] : u_l - u_{l'}(s_{l'}, v_{l'}) = t(v_{l'} - v_l) \text{ for some } t \in [0, T]\} \tag{4.43}$$

is of zero measure by a simple dimension argument. To see this, we first express

$u_{l'} = u_{\bar{l}} + s_{l'}(v_{\bar{l}} - v_{l'})$ by (3.1). Then the condition in (4.43) is

$$(t - s_{l'})v_{l'} = u_l - u_{\bar{l}} - s_{l'}v_{\bar{l}} + tv_l;$$

for given $t \neq s_{l'}$ the velocity $v_{l'}$ is contained in the countable set $\frac{1}{t-s_{l'}}(u_l - u_{\bar{l}} - s_{l'}v_{\bar{l}} + tv_l + \mathbb{Z}^d)$ giving restriction to a collection of $2 + (d - 1)$ dimensional sets. If on the other hand $t = s_{l'}$, then (4.43) gives $u_l - u_{\bar{l}} + t(-v_{\bar{l}} + v_l) = 0$. This implies that $v_{l'}$ is arbitrary but there is at most a finite number of $t \in [0, T]$ satisfying this equality in \mathbb{T}^d ; these subsets of $v_{l'}, v_{\bar{l}}, s_{l'}$ are hence $2d - 1$ dimensional.

Thus $\mathcal{MT}(\sigma, E) \setminus \mathcal{G}(0)$ is a countable union of manifolds of dimension less than or equal to $(2d)^{\#\sigma} - 1$. Since $\mathcal{MT} = \cup_{\sigma \in \mathcal{T}} \mathcal{MT}(\sigma)$ and $P_t \in L^1(\mathcal{MT})$ we obtain that $P_t(\mathcal{MT} \setminus \mathcal{G}(0)) = 0$ and $P_t(\mathcal{G}(0)) = 1$.

Furthermore, for each $\Phi \in \mathcal{G}(0)$ there exists a_0 such that $\Phi \in \mathcal{G}(a)$ for all $a < a_0$. Hence $\lim_{a \rightarrow 0} \mathcal{G}(a) = \mathcal{G}(0)$ and dominated convergence implies $\lim_{a \rightarrow 0} P_t(\mathcal{G}(a)) = P_t(\mathcal{G}(0)) = 1$. Thanks to the convergence in Proposition 4.6(ii) we obtain the remaining claim $\lim_{a \rightarrow 0} P_t^\alpha(\mathcal{G}(a)) = 1$. \square

We will also consider a finite number of trees simultaneously.

DEFINITION 4.14. For $\alpha > 1$ and $\Phi = (\Phi^{(1)}, \dots, \Phi^{(\alpha)}) \in \mathcal{MT}^\alpha$ we define

$$P_t^\alpha(\Phi) = \prod_{i=1}^\alpha P_t(\Phi^{(i)}),$$

and accordingly for $a > 0$.

The notion of good trees directly extends to a finite number of trees and the set of simultaneously good trees is denoted as $\mathcal{G}^\alpha(a) \subset \mathcal{MT}^\alpha$.

5. Empirical distribution. We consider the empirical distribution of trees \hat{P}_t defined by the Newton dynamics (2.3) and the rule that if the distance between two particles drops below a , then the trees of the particles are removed. The dependency on a will be suppressed throughout this section. We use the convention that $\text{Prob}_t(\cdot)$ denotes the joint distribution of n trees, whereas \hat{P}_t is the marginal distribution of an n -independent number of trees.

The choice of the initial states $z_1, \dots, z_n \in U \times \mathbb{R}^d$, time $t \geq 0$, and particle labels $(i_1, \dots, i_\alpha) \in \{1, \dots, n\}^\alpha$ induces a tree vector $(\Phi^{(i_1)}, \dots, \Phi^{(i_\alpha)}) \in \mathcal{MT}^\alpha$. We denote the induced joint probability of the trees by $\hat{P}_t^\alpha(\Phi^{(i_1)}, \dots, \Phi^{(i_\alpha)})$. Note that we can assume $i_1 = 1, i_2 = 2 \dots$ thanks to the invariance under permutation of the labels.

The key result in this section is the demonstration that the empirical distribution \hat{P}_t satisfies a differential equation (5.3) which is very similar to the idealized equation (4.5). The main difference is given by factors γ and ζ which account for dilution effects and initial overlaps. The similarity of (5.3) and (4.5) enables us to show later that the total variation distance between \hat{P}_t and P_t converges to 0 as a tends to 0.

The empirical Kolmogorov equation contains a singular gain term which is positive if and only if $t = \tau(\Phi)$ and a loss term which is nonzero for times $t > \tau(\Phi)$. Like in the idealized case the probability of a tree Φ is defined to be 0 if $t < \tau(\Phi)$. At time $t = \tau(\Phi)$ the probability $\hat{P}_t(\Phi)$ jumps from 0 to finite value. In contrast to the idealized case the collision probabilities depend on the structure of the tree, not only on the position and velocity of the root particle at time τ . However, due to the simplicity of the gainless evolution the collision rates can be expressed as functions of single-tree probabilities; hence we can avoid closure problems. This leads to the use of conditional probabilities as the particles which correspond to the nodes of the tree $\bar{\Phi}$ can have a direct influence on a further collision with the root of Φ' , e.g., by

ruling out particles that would contradict the data of $\bar{\Phi}$ and by reducing the number of available particles.

For $\Psi \in \mathcal{MT}^\alpha$ the conditional distribution of Φ given $\Psi \in \mathcal{MT}^\alpha$ is defined as

$$\hat{P}_t(\Omega | \Psi) = \text{Prob}_{a,t} \left(\Phi^{(1)} \in \Omega \mid (\Phi^{(2)}, \dots, \Phi^{(\alpha+1)}) = \Psi \right)$$

with the convention that $\hat{P}_t(\cdot | \Psi) = \hat{P}_t(\cdot)$ if $\Psi \in \mathcal{MT}^0$ or $\max\{\tau(\Psi^{(i)}) : i = 1 \dots \alpha\} \geq t$. We define the operator

$$\hat{Q}_t^a : PM(\mathcal{MT}) \times PM(\mathcal{MT}) \times \mathcal{MT} \rightarrow M(\mathcal{MT})$$

by

$$\hat{Q}_t^a[P, P'](\Phi) = P(\bar{\Phi}) \hat{L}_t^a[P'(\cdot | \bar{\Phi})](\Phi),$$

where

$$(5.1) \quad \hat{L}_t^a[P](\Phi) = \delta(t - \tau(\Phi)) P(\Phi') [(v(\Phi') - v(\Phi)) \cdot \nu]_+ - \hat{\mu}_t[P](\Phi).$$

The empirical collision rate $\hat{\mu}_t$ is obtained by considering all possible colliding initial data; we define it by the expression

$$\hat{\mu}_t[P](\Phi) = \frac{1}{\int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_\Phi(z_t)} \int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_\Phi(z_t) [(v - v') \cdot \nu]_+,$$

$$\mathbf{1}_\Phi(u', v') = \begin{cases} 1 & \text{if } \min\{|u' - u_l + s(v' - v_l)| : s \in [0, s_l], l \in m(\Phi) \setminus \text{root}\} > a, \\ 0 & \text{else} \end{cases}$$

with the convention

$$(5.2) \quad z_t = (u + t(v - v') + av, v).$$

PROPOSITION 5.1. *Let $\alpha \geq 0$, $\Gamma \in \mathcal{MT}^\alpha$ and a sufficiently small. The empirical distribution \hat{P}_t satisfies for all $(\Phi, \Gamma) \in \mathcal{G}^{1+\alpha}(a)$ the following differential equation:*

$$(5.3) \quad \partial_t \hat{P}_t(\Phi | \Gamma) = (1 - \gamma) \hat{Q}_t^a[\hat{P}_t(\cdot | \Gamma), \hat{P}_t(\cdot | \Gamma)](\Phi),$$

$$(5.4) \quad \hat{P}_0(\Phi | \Gamma) = \zeta(\Phi, \Gamma) f_0(u(\Phi), v(\Phi)) \mathbf{1}_{\#m(\Phi)=1},$$

where $\gamma = (\#m(\bar{\Phi}) + \#m(\Gamma))a^{d-1}$ and

$$(5.5) \quad \zeta(\Phi, \Gamma) = \left(1 - \frac{\int_{\{(u',v) : |u'-u(\Phi)| \leq a\}} df_0(u', v)}{\int_{\{(u',v) : \min\{|u'-u_l| : l \in m(\Gamma)\} \geq a\}} df_0(u', v)} \right)^{n - \#m(\Gamma) - 1} \in [0, 1].$$

The proof of Proposition 5.1 relies on the a priori information that \hat{P}_t is absolutely continuous with respect to the Lebesgue measure.

LEMMA 5.2. *Let $a > 0$ and $\Psi \in \mathcal{G}(a)$. The empirical distribution \hat{P}_t is absolutely continuous with respect to the Lebesgue measure λ in a neighborhood of Ψ . Moreover, if $A \subset U \times \mathbb{R}^d$ is measurable and has the property $\mathbf{1}_\Psi(z) = 1$ for all $z \in A$, then the inequality*

$$(5.6) \quad \hat{P}_t \left(z_{\text{root}}^{(1)}, z_{\text{root}}^{(2)} \in A \mid \Psi \right) \leq \left(\frac{\int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_A(z_t)}{\int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_\Psi(z_t)} \right)^2$$

holds with z_t given by (5.2).

Proof. Note first that $\mathcal{G}(a)$ is an open set. This implies that there exists $h > 0$ such that $B_h(\Psi) \subset \mathcal{G}(a)$. If $\#m(\Psi) = 1$, then $\hat{P}_t(\Phi) \leq f_0(z_0)$ for all $t \geq 0$. This establishes the absolute continuity.

Assume now that $\#m(\Psi) > 1$ and let $\varphi : B_h(\Psi) \rightarrow \mathcal{MT} \times \mathcal{MT}$ be defined by $\varphi(\Psi) = (\bar{\Psi}, \Psi')$. Then

$$(5.7) \quad \det(\nabla\varphi) = a^{d-1}(v - v') \cdot \nu.$$

Indeed, assume that $j \in m(\Phi)$ is the node which corresponds to the last collision (i.e., $s_j = \tau(\Phi)$). A simple algebraic computation shows that in a coordinate system where $\nu_j = e_1 = (1, 0, \dots, 0)^T$ we obtain that the gradient $F_{lk} = \nabla_{t_l, \nu_l} \varphi_k(\Phi) \in \mathbb{R}^{d \times d}$ if $l \neq \text{root}$ and $F_{lk} = \nabla_{u_l} \varphi_k(\Phi)$ if $l = \text{root}$. Then one obtains, in a coordinate system where $\nu_\lambda = e_1 = (1, 0, \dots, 0)^T$, the formula

$$(5.8) \quad F_{lk} = \begin{cases} \begin{pmatrix} (v_l - v_{\bar{l}}) \cdot \nu_l & 0 \\ & a \text{Id}(d-1) \end{pmatrix} & \text{if } k = l = j, \\ \text{Id}(d) & \text{if } l = k \neq j, \\ \text{Id}(d) & \text{if } l = \text{root}, k = j, \\ 0 & \text{else,} \end{cases}$$

where $\text{Id}(d)$ is the d -dimensional identity matrix. Hence, the determinant of F is just the product of the determinants of F_{ll} . This yields (5.7).

Now we consider the map $\tilde{\varphi} : B_h(\Psi) \rightarrow (U \times \mathbb{R}^d)^{\#m(\Psi)}$ which assigns to each node l the initial values $z_l \in U \times \mathbb{R}^d$. Iterating formula (5.7) implies

$$(5.9) \quad \det(\nabla\tilde{\varphi}(\Psi)) = \prod_{l \in m(\Psi) \setminus \text{root}} [a^{d-1}(v_l - v_{\bar{l}}) \cdot \nu_l].$$

Observe next that the inequality

$$\hat{P}_t(\Psi) \leq \text{Prob}_a(\{\tilde{\varphi}_l(\Psi) : l \in m(\Psi)\} \subset \omega)$$

holds because the existence of a tree requires that the initial states form a subset of ω .

Let $\omega \subset U \times \mathbb{R}^d$ be the set of initial positions and velocities. Now we consider the cubes $C_{h,l} \subset U \times \mathbb{R}^d$ centered at $\tilde{\varphi}_l(\Psi)$ with side length h . The cubes are disjoint for sufficiently small h since $\Psi \in \mathcal{G}$. If a and h are small, the assumption that the initial values $z_1 \dots z_n \in U \times \mathbb{R}^d$ are iid random variables with law f_0 and the scaling law (1.1) imply that

$$(5.10) \quad \text{Prob}_a(\#\omega \cap C_{h,l} = 1) = a^{1-d} f_0(C_{h,l})(1 + o(1)) \text{ as } h \rightarrow 0$$

for each $l \in m(\Psi)$, where by a slight abuse of notation we use f_0 as a measure, i.e., $f_0(C_{h,l}) = \int_{C_{h,l}} f_0(u, v) \, du \, dv$. Let now $C_h(\Psi) = \prod_{l \in m(\Psi)} C_{h,l}(\Psi)$. The formula

$$\lambda(\tilde{\varphi}^{-1}(C_h)) = h^{2d\#m(\Psi)} \det(F(\Psi))^{-1}(1 + o(1)) \text{ as } h \rightarrow 0$$

together with (5.9) implies that

$$(5.11) \quad \frac{\hat{P}_t(\tilde{\varphi}^{-1}(C_h))}{\lambda(\tilde{\varphi}^{-1}(C_h))} = \frac{f_0(C_{h,\text{root}})}{h^{2d}} \prod_{l \in m(\Psi) \setminus \text{root}} \left[\frac{f_0(C_{h,l})}{h^{2d}(1 + o(1))} (v_l - v_{\bar{l}}) \cdot \nu_l \right].$$

Since $f_0 \in L^1(U \times \mathbb{R}^d)$, the right-hand side in (5.11) remains bounded as $h \rightarrow 0$ for almost every $\Psi \in \mathcal{G}(a)$. This establishes the absolute continuity of \hat{P}_t . Formula (5.6)

is an immediate consequence of the definition of $\hat{P}(\cdot | \Psi)$ and the assumption that z_1, z_2, \dots are iid random variables with law f_0 . We obtain an upper bound instead of equality because of the possibility that particles 1 and 2 could correspond to one of the nodes of Ψ . The proof of the lemma is complete. \square

Proof of Proposition 5.1. To simplify the notation we assume that $\Gamma = \emptyset$, which means that no conditioning is active. The general case is analogous, as explained at the end of the proof.

We consider for a fixed $t \geq 0$ two cases: $\tau(\Phi) = t$ and $\tau(\Phi) < t$. We start with the case $\tau(\Phi) = t$ and will derive the product expression in the singular part of \hat{Q}_t . Recall that $\nu(\Phi)$ denotes the impact parameter of the final collision.

The probability $\hat{P}_t(\Psi)$ can be expressed in terms of the probabilities of Ψ'_t and $\bar{\Psi}_t$:

$$\begin{aligned} \hat{P}_t(\Psi) &= \hat{P}_t(\Phi^{(1)} = \Psi) = \hat{P}_t(\Phi^{(1)'} = \Psi'_t \text{ and } \bar{\Phi}^{(1)} = \bar{\Psi}_t) \\ &= \hat{P}_t(\Phi^{(1)'} = \Psi'_t \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t) \hat{P}_t(\bar{\Phi}^{(1)} = \bar{\Psi}_t). \end{aligned}$$

The key idea is that the first probability can be expressed in terms of a two-body event. We will now demonstrate that

$$\hat{P}_t(\Phi^{(1)'} = \Psi'_t \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t) = (1 - \gamma) \hat{P}_t(\Phi^{(1)} = \Psi'_t \mid \Phi^{(2)} = \bar{\Psi}_t) [(v - v') \cdot \nu]_+ \tag{5.12}$$

by establishing matching upper and lower bounds.

First we derive the upper bound. For each $t \geq 0$ and $h > 0$ sufficiently small define the set of trees near Ψ , which have identical pruned trees

$$U_h(\Psi) = \{\Phi \in B_h(\Psi) : \Phi_l = \Psi_l \text{ for all } l \in m(\bar{\Psi}_t)\}.$$

From this we introduce

$$V_h = \{\Phi' : \Phi \in U_h\},$$

which are all possible nearby extracted trees. As \hat{P}_t is absolutely continuous by Lemma 5.2 we obtain that

$$\hat{P}_t(\Psi'_t \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t) = \lim_{h \rightarrow 0} h^{-2d\#m(\Psi')} \hat{P}_t(\Phi^{(1)} \in U_h \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t).$$

Since there are at most n possible choices for the index of the colliding particle we find that

$$\hat{P}_t(\Phi^{(1)} \in U_h \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t) \leq \sum_{i=1}^n \hat{P}_t(\Phi^{(i)} \in V_h \mid \Phi^{(1)} = \bar{\Psi}_t). \tag{5.13}$$

The permutation invariance and the fact that the particles in $\bar{\Psi}_t$ are ruled out as collision partners implies that

$$\sum_{i=1}^n \hat{P}_t(\Phi^{(i)} \in V_h \mid \Phi^{(1)} = \bar{\Psi}_t) = (n - \#m(\bar{\Psi}_t)) \hat{P}_t(\Phi^{(1)} \in V_h \mid \Phi^{(2)} = \bar{\Psi}_t).$$

Formula (5.7) implies that

$$\begin{aligned} &\lim_{h \rightarrow 0} h^{-2d\#m(\Psi')} \hat{P}_t(\Phi^{(1)} \in V_h \mid \Phi^{(2)} = \bar{\Psi}_t) \\ &= \hat{P}_t(\Phi^{(1)} = \Psi'_t \mid \Phi^{(2)} = \bar{\Psi}_t) a^{d-1} [(v - v') \cdot \nu]_+ \end{aligned}$$

and thereby delivers the upper bound

$$(5.14) \quad \hat{P}_t \left(\Phi^{(1)} = \Psi'_t \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t \right) \leq (1 - \gamma) \hat{P}_t \left(\Phi^{(1)} = \Psi'_t \mid \Phi^{(2)} = \bar{\Psi}_t \right) [(v - v') \cdot \nu]_+.$$

Next we derive the corresponding lower bound

$$(5.15) \quad \hat{P}_t \left(\Phi_t^{(1)'} = \Psi'_t \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t \right) \geq (1 - \gamma) \hat{P}_t \left(\Phi^{(1)} = \Psi'_t \mid \Phi^{(2)} = \bar{\Psi}_t \right) [(v - v') \cdot \nu]_+.$$

Define the set of initial values leading to a collision within the time interval $[t, t + h]$:

$$(5.16) \quad W_h(t) = \left\{ (u', v') \in U \times \mathbb{R}^d \mid \exists (\nu, t') \in S^{d-1} \times \mathbb{R} \text{ such that } \right. \\ \left. u - u' + t'(v - v') + a\nu = 0 \text{ and } (v - v') \cdot \nu \geq 0 \text{ and } t \leq t' \leq t + h \right\}.$$

Clearly

$$\hat{P}_t \left(\Phi^{(1)} \in U_h \mid \bar{\Phi}_t^{(1)} = \bar{\Psi}_t \right) \geq \sum_{i=1}^n \hat{P}_t \left(\Phi^{(i)} \in V_h \mid \bar{\Phi}_t^{(1)} = \bar{\Psi}_t \right) \\ - \text{Prob} \left(\#(\omega \cap W_h) \geq 2 \mid \bar{\Phi}^{(1)} = \bar{\Psi}_t \right),$$

and using the inclusion-exclusion principle we obtain that

$$(5.17) \quad \text{Prob} \left(\#(\omega \subset W_h) \geq 2 \mid \bar{\Phi}_t^{(1)} = \bar{\Psi}_t \right) \leq \sum_{1 \leq i < j \leq n} \hat{P}_t \left(\{z^{(i)}, z^{(j)}\} \subset W_h \mid \bar{\Phi}_t^{(1)} = \bar{\Psi}_t \right).$$

Lemma 5.2 implies that

$$(5.18) \quad \hat{P}_t \left(\{z^{(1)}, z^{(2)}\} \subset W_h \mid \bar{\Phi}_t^{(3)} = \bar{\Psi}_t \right) \leq (I_h)^2$$

with

$$I_h = \frac{\int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_{W_h}(z_t) \mathbf{1}_{\Psi}(z_t)}{\int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_{\Psi}(z_t)}.$$

The estimation of I_h is straightforward: formula (5.8) implies that

$$(5.19) \quad \int f_0(u, v) \mathbf{1}_{V_h}(u, v) dudv = a^{d-1} h^{2d} (f(u', v') [(v - v') \cdot \nu]_+ + o(1))$$

as h tends to 0, where u', v' are the root data of Ψ' . Moreover, since $\Psi \in \mathcal{G}(a)$ one obtains that there exists a constant C uniformly on $\mathcal{G}(a)$ such that

$$(5.20) \quad \int f_0(u, v) \mathbf{1}_{\Psi}(u, v) du dv \geq 1 - \kappa_d C K_\infty a^{d-1} \geq \frac{1}{2}$$

if a is sufficiently small. As a consequence of the bounds (5.19) and (5.20) the right-hand side in estimate (5.18) tends to 0 as h tends to 0. Hence, we have established that the left-hand side of (5.17) converges to 0 and thus (5.15) holds.

Combining (5.14) and (5.15) one obtains (5.3) in the case that $\tau = t$.

The final step is the justification of the loss term. It is not possible to obtain the loss term by integrating the gain term for two reasons:

1. The explicit representation of the gain term is valid only for good trees.
2. The representation of loss term is simpler than the conservation form because of cancellation effects.

Assume next that $\tau(\Phi) < t$ and consider the set of colliding initial values W_h defined in (5.16). It suffices to show that

$$(5.21) \quad \lim_{h \rightarrow 0} \frac{1}{h} (\hat{P}_{t+h}(\Phi) - \hat{P}_t(\Phi)) = - \lim_{h \rightarrow 0} \frac{1}{h} \hat{P}_t(\#(\omega \cap W_h(t)) > 0 \mid \Psi) = -(1 - \gamma)J$$

with

$$J = \frac{\int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_\Psi(z_t) [(v - v') \cdot \nu]_+}{\int_{S^{d-1}} d\nu \int_{\mathbb{R}^d} dv' f_0(z_t) \mathbf{1}_\Psi(z_t)}.$$

Analogously to the case $\tau(\Phi) = t$ we obtain the upper and lower bounds

$$0 \leq (1 - \gamma)J - \frac{1}{h} \hat{P}_t(\#(\omega \cap W_h(t)) > 0 \mid \Psi) = o(1)$$

as $h \rightarrow 0$. Equation (5.21) holds thanks to (5.20) and the fact that J does not depend on h . Thus we have established the claim also in the case $\tau < t$.

Finally we establish the initial condition (5.4). If $\Gamma = \emptyset$ and u_1 has been selected, the probability that each of the remaining $n - 1$ initial positions does not overlap is given by (5.5).

The assumption that $\Gamma = \emptyset$ does not involve a loss of generality. For generic Γ everything can be repeated line by line, except that we have to add the conditioning encoded by Γ to all expressions involving \hat{P} . For the initial conditions (5.4), we need to condition on the event that the remaining initial positions also do not overlap with the particle in Γ . The expression (5.5) for $\zeta(\Phi, \Gamma)$ follows when observing that the pair of trees (Φ, Γ) are assumed to be good such that the particles of Φ and Γ do not overlap for a small enough. \square

5.1. Convergence. We now proceed to estimate the difference between the empirical distribution \hat{P}_t^a and the idealized P_t . The key estimate which provides a quantitative link between \mathcal{Q} and $\hat{\mathcal{Q}}$ is provided by the following comparison principle.

PROPOSITION 5.3. *Let $\Phi \in \mathcal{MT}$, $\Psi \in \mathcal{MT}^\alpha$ for some $\alpha > 0$ such that $(\Phi, \Psi) \in \mathcal{G}^{\alpha+1}$. Then the estimate*

$$(5.22) \quad 1 - \frac{\hat{P}_t(\Phi \mid \Psi)}{\zeta(\Phi, \Psi) P_t(\Phi)} \leq \rho_t(\#m(\Phi))\eta_t$$

holds for all $\Phi \in \mathcal{G}$ with ζ defined in (5.5), $\rho_t(k) = (2k - 1)Ct \exp(Ct)$,

$$\eta_t(\Phi, \Psi) = \int f_0(u, v) (1 - \mathbf{1}_\Phi(u, v) \mathbf{1}_\Psi(u, v))(1 + |v|) du dv,$$

$$C(\Phi) = 2 \max \left\{ \sup \left\{ \mu_t(u, v) : u \in U, |v| \leq \max_{l \in m(\Phi)} |v_l| \right\}, \right. \\ \left. \max \{|v_l - v_{l'}| : l, l' \in m(\Phi)\} \right\},$$

and μ_t is defined in (4.16).

Note that only η_t is affected by the conditioning.

Proof. To simplify the notation we define $p_t^a(\Phi) = \zeta(\Phi, \Psi) P_t^a(\Phi)$. Observe that ρ_t is superadditive, i.e., $\rho_t(x + y) \geq \rho_t(x) + \rho_t(y)$. We use induction over α . The case $\alpha = 1$ will be treated below. If $\alpha > 1$, then we can split Φ into two trees $\Phi^* \in \mathcal{MT}^1$ and $\Phi^{**} \in \mathcal{MT}^{\alpha-1}$. Clearly

$$\begin{aligned} \hat{P}_t(\Phi | \Psi) &= \hat{P}_t(\Phi^* | \Phi^{**}, \Psi) \hat{P}_t(\Phi^{**} | \Psi) \\ &\geq (1 - \rho_t(\#m(\Phi^*))\eta_t)(1 - \rho_t(\#m(\Phi^{**}))\eta_t) p_t^a(\Phi^* | \Phi^{**}) p_t^a(\Phi^{**}) \\ &\geq (1 - (\rho_t(\#m(\Phi^*)) + \rho_t(\#m(\Phi^{**})))\eta_t) p_t^a(\Phi) \geq (1 - \rho_t(\#m(\Phi))\eta_t) p_t^a(\Phi). \end{aligned}$$

The first inequality holds because of the induction assumption, the second one is due to the sign of the quadratic term, and the third inequality is a consequence of the superadditivity of the function ρ_t . Thus, it suffices to consider the case $\alpha = 1$.

First, note that the definitions of L_t, \hat{L}_t in (4.11) and (5.1) imply that for every probability measure P with marginal $\mu_t(u, v) = f_0(u - tv, v)$ the inequality

$$(5.23) \quad (1 - \gamma) \hat{L}_t[P](\Phi) \geq L_t^a[P](\Phi) \begin{cases} 1 + 2\eta & \text{if } \tau(\Phi) \neq t, \\ 1 - \gamma & \text{if } \tau(\Phi) = t \end{cases}$$

holds, where $\tau > 0$ is the time of the final collision of Φ . If $t > \tau$, then the first term in (5.23) is relevant and one obtains

$$\begin{aligned} (5.24) \quad \frac{\partial}{\partial t}(\hat{P}_t - p_t^a)(\Phi) &= (1 - \gamma) \hat{L}_t[\hat{P}_t](\Phi) \hat{P}_t(\Phi) - L_t^a[P_t^a](\Phi) p_t^a(\Phi) \\ &\geq (1 + 2\eta) L_t^a[\hat{P}_t](\Phi) \hat{P}_t(\Phi) - L_t^a[P_t^a](\Phi) p_t^a(\Phi) \\ &= (1 + 2\eta) L_t^a[\hat{P}_t](\Phi) \left(\hat{P}_t(\Phi) - p_t^a(\Phi) \right) \\ &\quad + (1 + 2\eta) (L_t^a[\hat{P}_t](\Phi) - L_t^a[P_t^a](\Phi)) p_t^a(\Phi) + 2\eta L_t^a[P_t^a](\Phi) p_t^a(\Phi). \end{aligned}$$

For fixed Φ this is a one-dimensional ordinary differential inequality for $x(t) = (\hat{P}_t - p_t^a)(\Phi)$ and delivers an integrated estimate. For the one-dimensional ODE, if $\dot{x}(t) \geq a(t)x(t) + b(t)$ and $x(\tau) = x_0$, i.e.,

$$\begin{aligned} \dot{x}(t) &= a(t)x(t) + b(t) + c(t) \text{ with } c(t) \geq 0, x(\tau) = x_0 \\ \dot{y}(t) &= a(t)y(t) + b(t) \text{ with } y(\tau) = x_0, \end{aligned}$$

then by the variation of constants formula we have

$$\begin{aligned} x(t) &= \exp\left(\int_{\tau}^t a(s) ds\right) x_0 + \int_{\tau}^t \exp\left(\int_s^t a(\sigma) d\sigma\right) (b(s) + c(s)) ds \\ &\geq y(t) = \exp\left(\int_{\tau}^t a(s) ds\right) x_0 + \int_{\tau}^t \exp\left(\int_s^t a(\sigma) d\sigma\right) b(s) ds \end{aligned}$$

for $t \geq \tau$, independent of the signs of a and b . Then estimating

$$|\hat{\mu}_s - \mu_s|_{L^\infty(U \times \mathbb{R}^d)} \leq 2\eta s$$

and observing that $s \mapsto \eta_s$ is nondecreasing we obtain

$$\begin{aligned} & \hat{P}_t(\Phi) - p_t^a(\Phi) \\ & \geq \exp\left((1 + 2\eta_t) \int_{\tau}^t L_s^a[\hat{P}_s] ds\right) \left(\hat{P}_{\tau}(\Phi) - p_{\tau}^a(\Phi)\right) \\ & \quad + \int_{\tau}^t (1 + 2\eta_t) 2\eta_t \exp\left((1 + 2\eta_t) \int_s^t L_{s'}^a[\hat{P}_{s'}](\Phi) ds'\right) p_s^a(\Phi) ds \\ & \quad + 2\eta_t \int_{\tau}^t \exp\left((1 + 2\eta_t) \int_s^t L_{s'}^a[\hat{P}_{s'}](\Phi) ds'\right) L_s^a[P_s^a](\Phi) p_s^a(\Phi) ds. \end{aligned}$$

After observing that

$$\exp\left(\int_s^t L_{s'}^a[P_{s'}^a] ds'\right) p_s^a(\Phi) = p_t^a(\Phi)$$

we obtain

$$\begin{aligned} \hat{P}_t(\Phi) - p_t^a(\Phi) & \geq \exp\left((1 + 2\eta_t)\eta_t \int_{\tau}^t L_s^a[\hat{P}_s] ds\right) \left(\hat{P}_{\tau}(\Phi) - p_{\tau}^a(\Phi)\right) \\ & \quad + \left[\int_{\tau}^t (1 + 2\eta_t) 2\eta_t \exp\left((1 + 2\eta_t) \int_s^t L_{s'}^a[\hat{P}_{s'}](\Phi) ds'\right) \right. \\ & \quad \quad \left. \exp\left(- \int_s^t L_{s'}^a[P_{s'}^a](\Phi) ds'\right) ds \right. \\ & \quad \quad + 2\eta_t \int_{\tau}^t \exp\left((1 + 2\eta_t) \int_s^t L_{s'}^a[\hat{P}_{s'}] ds'\right) L_s^a[P_s^a](\Phi) \\ (5.25) \quad & \quad \left. \exp\left(- \int_s^t L_{s'}^a[P_{s'}^a](\Phi) ds'\right) ds \right] p_t^a(\Phi). \end{aligned}$$

We use induction over $k = \#m(\Phi)$ in (5.5). First assume that $\#m(\Phi) = 1$. In this case $\tau = 0$ and $\hat{P}_0(\Phi) = p_0^a(\Phi)$, which together with (5.25) implies that

$$\begin{aligned} & \hat{P}_t(\Phi) - p_t^a(\Phi) \\ & \geq \left[\int_{\tau}^t (1 + 2\eta_t) 2\eta_t \exp\left((1 + 2\eta_t) \int_s^t L_{s'}^a[\hat{P}_{s'}] ds'\right) \exp\left(- \int_s^t L_{s'}^a[P_{s'}^a](\Phi) ds'\right) ds \right. \\ & \quad + 2\eta_t \int_{\tau}^t \exp\left((1 + 2\eta_t) \int_s^t L_{s'}^a[\hat{P}_{s'}] ds'\right) L_s^a[P_s^a](\Phi) \\ & \quad \left. \exp\left(- \int_s^t L_{s'}^a[P_{s'}^a](\Phi) ds'\right) ds \right] p_t^a(\Phi) \\ & \geq -Ct \exp(Ct) \eta_t p_t^a(\Phi). \end{aligned}$$

Assume next that the estimate has been established for all trees with at most k nodes and let $\#m(\Phi) = k + 1$. Define $k_1 = \#m(\bar{\Phi})$ and $k_2 = \#m(\Phi')$ with $k_1 + k_2 = k + 1$ and $\max\{k_1, k_2\} \leq k$. Thus the induction assumption implies that

$$\begin{aligned} \hat{P}_{\tau}(\bar{\Phi}) & \geq (1 - \rho_{\tau}(k_1)) p_{\tau}^a(\bar{\Phi}), \\ \hat{P}_{\tau}(\Phi' \mid \bar{\Phi}) & \geq (1 - \rho_{\tau}(k_2)) p_{\tau}^a(\Phi'). \end{aligned}$$

Using (5.25) and $p_\tau^a(\bar{\Phi})p_\tau^a(\Phi') \leq Cp_\tau^a(\Phi)$ from the definition of P_t^a we find that

$$\begin{aligned} & \hat{P}_t(\Phi) - p_t^a(\Phi) \\ & \geq \exp\left((1 + 2\eta) \int_\tau^t L_s^a[\hat{P}_s] ds \right) [(v - v') \cdot \nu]_+ \\ & \quad \times \left(\underbrace{\hat{P}_\tau(\bar{\Phi})\hat{P}_\tau(\Phi' | \bar{\Phi})}_{\geq p_\tau^a(\bar{\Phi})p_\tau^a(\Phi')(1 - (\rho_t(k_1) + \rho_t(k_2)))} - p_\tau^a(\bar{\Phi})p_\tau^a(\Phi') \right) \\ & \quad + \left[\int_\tau^t (1 + 2\eta)2\eta \exp\left((1 + 2\eta) \int_s^t L_{s'}^a[\hat{P}_{s'}](\Phi) ds' \right) \exp\left(- \int_s^t L_{s'}^a[P_{s'}^a](\Phi) ds' \right) ds \right. \\ & \quad \left. + 2\eta \int_\tau^t \exp\left((1 + 2\eta) \int_s^t L_{s'}^a[\hat{P}_{s'}] ds' \right) L_s^a[P_s^a](\Phi) \right. \\ & \quad \left. \exp\left(- \int_s^t L_{s'}^a[P_{s'}^a](\Phi) ds' \right) ds \right] p_t^a \\ & \geq -(\rho_\tau(k_1) + \rho_\tau(k_2)) p_\tau^a(\bar{\Phi})p_\tau^a(\Phi') [(v - v') \cdot \nu]_+ + Ct \exp(Ct)\eta p_t^a(\Phi) \\ & \geq -(\rho_t(k_1) + \rho_t(k_2)) p_t^a(\bar{\Phi}) + Ct \exp(Ct)\eta p_t^a(\Phi). \end{aligned}$$

Since

$$(\rho_t(k_1) + \rho_t(k_2)) + Ct \exp(Ct) \leq Ct \exp(Ct)(2k_1 + 2k_2 - 2 + 1) \leq \rho_t(k + 1),$$

inequality (5.22) has been established. \square

Recall the function $M(a)$ in Definition 4.12 and the constant ζ .

LEMMA 5.4. *There exists a constant $C > 0$ which depends on K_∞ and α such that*

$$\inf_{\Phi, \Gamma \in \mathcal{G}(a)} \zeta(\Phi, \Gamma) \geq 1 - Ca(1 + a^d M(a)).$$

Proof. Formulas (5.5), (1.6), and (1.1) imply

$$\zeta(\Phi, \Gamma) \geq 1 - 2n \frac{K_\infty \kappa_{d+1} a^d}{1 - K_\infty \#m(\Gamma) \kappa_{d+1} a^d} \geq 1 - Ca(1 + a^d M(a))$$

if C is suitably chosen and a is sufficiently small. \square

Estimate (5.22) immediately implies that \hat{P}_t converges in the total variation sense to P_t as a tends to 0.

PROPOSITION 5.5. *For each $\alpha \geq 1$ the total variation distance between P^α and \hat{P}^α satisfies*

$$\lim_{a \rightarrow 0} \|P_t^{\alpha, \alpha} - \hat{P}_t^\alpha\|_{L^1(\mathcal{MT})} = 0.$$

Proof. We assume that $\alpha = 1$; the case $\alpha > 1$ can be treated analogously. Lemma 4.13 implies that for each $\varepsilon > 0$ there exists $a > 0$ so small that

$$P_t^a(\mathcal{MT} \setminus \mathcal{G}(a)) \leq \varepsilon.$$

We use that

$$\|P_t^a - \hat{P}_t\|_{TV} = 2 \sup_{\Omega} (P_t^a(\Omega) - \hat{P}_t(\Omega)).$$

For each $\Omega \subset \mathcal{MT}$ Lemma 5.4 implies

$$\begin{aligned} & P_t^a(\Omega) - \hat{P}_t(\Omega) \\ &= P_t^a(\Omega \cap \mathcal{G}(a)) - \hat{P}_t(\Omega \cap \mathcal{G}(a)) + P_t^a(\Omega \setminus \mathcal{G}(a)) - \hat{P}_t(\Omega \setminus \mathcal{G}(a)) \\ (5.26) \quad &\leq (\rho\xi(a) + Ca)P_t^a(\mathcal{MT}) + P_t^a(\mathcal{MT} \setminus \mathcal{G}(a)) \leq (\rho\xi + Ca) + P_t^a(\mathcal{MT} \setminus \mathcal{G}(a)), \end{aligned}$$

where $\xi(a) := \sup_{\Phi \in \mathcal{G}(a)} \eta_t(\Phi)$ with ρ_t and $\eta_t(\Phi)$ defined in Proposition 5.3.

Uniformly for $\Phi \in \mathcal{G}(a)$ a particle can cover a cylinder of volume less equal to $V(a)t\kappa_d a^{d-1}$. Assume now that $M(a), V(a) \leq a^{-\frac{1}{4}}$ and note that $\#m(\Phi) \leq M(a)$ and that the u -marginal of f_0 is bounded in $L^\infty(U)$ by K_∞ . Then we obtain

$$(5.27) \quad \xi(a) \leq 2M(a)V(a)t\kappa_d a^{d-1}K_\infty \leq Ca^{d-3/2}$$

and $\rho_t \leq Ca^{-1/4}$ for some constant $C > 0$. Then $\lim_{a \rightarrow 0} \xi(a)\rho(a) = 0$ and by Lemma 4.13 $\lim_{a \rightarrow 0} P_t^a(\mathcal{MT} \setminus \mathcal{G}(a)) = 0$, which implies the claim. \square

6. Effective dynamics.

Proof of Theorem 2.1. We first show that the distribution of a single tagged particle satisfies the gainless Boltzmann equation (2.5). Let $A \subset U \times \mathbb{R}^d$ and define $\Omega_t(A) \subset \mathcal{MT}$ by

$$(6.1) \quad \Omega_t(A) = \{\Phi : \beta(\Phi) = 1 \text{ and } (u + tv, v) \in A\}.$$

According to Proposition 4.10 every weak solution f_t of (4.34) is a mild solution and thereby unique. Proposition 4.9 implies that

$$\int_A df_t = P_t(\Omega_t),$$

and thus

$$\left| \lim_{a \rightarrow 0} \hat{P}_t(\Omega_t) - \int_A df(u, v, t) \right| \stackrel{\text{Prop. 4.9}}{=} \lim_{a \rightarrow 0} \left| \hat{P}_t(\Omega_t) - P_t(\Omega_t) \right| \stackrel{\text{Prop. 5.5}}{=} 0.$$

The convergence is uniform in A by (5.26).

Next we consider the random variables for $i \in \{1, \dots, n\}$

$$\chi_i(t) = \begin{cases} 1 & \text{if } (u_i(t), v_i(t)) \in A \text{ and } \beta_i^{(a)}(t) = 1, \\ 0 & \text{else.} \end{cases}$$

Then the first part of the proof yields that $\lim_{a \rightarrow 0} \langle \chi_i(t) \rangle = \int_A df_i(u, v)$ for each i . Now we define the random variable $s_n = \frac{1}{n} \sum_{i=1}^n \chi_i(t)$. The claim (2.9) follows if the variance $V_n = \langle (s_n - \langle s_n \rangle)^2 \rangle$ converges to 0 as n tends to infinity. Thanks to the permutation invariance we obtain that

$$\begin{aligned} V_n &\leq \frac{1}{n} (\chi_1(t) - \langle \chi_1(t) \rangle)^2 + \frac{n-1}{n} \langle (\chi_1(t) - \langle \chi_1(t) \rangle)(\chi_2(t) - \langle \chi_2(t) \rangle) \rangle \\ &\leq \frac{1}{n} + |\langle (\chi_1(t) - \langle \chi_1(t) \rangle)(\chi_2(t) - \langle \chi_2(t) \rangle) \rangle|. \end{aligned}$$

If we apply Proposition 5.5 again with $\alpha = 2$ and

$$\Omega_t(A, B) = \{ \Phi \in \mathcal{MT}^2 \mid \beta_1(\Phi) = \beta_2(\Phi) = 1, (u_1 + tv_1, v_1) \in A \text{ and } (u_2 + tv_2, v_2) \in B \}$$

we obtain that

$$\lim_{a \rightarrow 0} \langle \chi_1(t) \chi_2(t) \rangle = \lim_{a \rightarrow 0} \hat{P}_t^2(\Omega(A, A)) = P_t(\Omega(A))^2.$$

This implies that

$$(6.2) \quad \lim_{a \rightarrow 0} \langle (\chi_1(t) - \langle \chi_1(t) \rangle)(\chi_2(t) - \langle \chi_2(t) \rangle) \rangle = 0$$

uniformly in A , which completes the proof of (2.9). Equation (6.2) is the main reason to consider trees with several roots. In particular this gives $V_n \leq b(n)$ for some decaying function $b : \mathbb{N} \rightarrow \mathbb{R}$ uniformly in the test set A again by (5.26).

Finally we show (2.10). We recall a well-known principle in probability theory. Let $x_N \in \mathbb{R}$ be a sequence of independent random numbers such that $\mathbb{E}(x_N) = 0$ and let V_N be the variance of x_N . If $\sum_{N=1}^\infty V_N < \infty$, then almost surely $\lim_{N \rightarrow \infty} x_N = 0$.

Indeed, for every $\varepsilon, N_0 > 0$ Tchebychev’s inequality yields the estimate

$$\text{Prob} \left(\sup_{N \geq N_0} |x_N| \leq \varepsilon \right) \geq \prod_{N=N_0}^\infty \left(1 - \frac{V_N}{\varepsilon^2} \right) \geq 1 - \frac{1}{\varepsilon^2} \sum_{N=N_0}^\infty V_N.$$

Hence $\lim_{a \rightarrow 0} \text{Prob}(\sup_{N \geq N_0} |x_N| \leq \varepsilon) = 1$, i.e., for each realization and each $\varepsilon > 0$ there exists almost surely a number $N_0 > 0$ such that $\sup_{N \geq N_0} |x_N| \leq \varepsilon$.

Let s_n be the sum and V_n be the variance of s_n as above. Since $\lim_{a \rightarrow 0} V_n = 0$ uniformly in A there exists a subsequence V_{n_k} such that $\sum_{k=1}^\infty V_{n_k} < \infty$ for all A . We apply now the previous consideration to the sequence $x_k = s_{n_k}$ such that

$$\int_A \frac{1}{n_k} \sum_{i=1}^{n_k} \beta_i^{(a_k)}(t) \delta(\cdot - (u_i(t), v_i(t))) \, du \, dv \rightarrow \int_A df_t$$

as k tends to infinity and thus we obtain the desired weak- $*$ convergence (2.10). □

7. Spatial concentrations. We discuss variants and limitations of the presented theory. We require $\int_{\mathbb{R}^d} f_0(\cdot, v) \, dv \in L^\infty(U)$ in (1.6). This implies that in (5.27) the expected number of particles overlapping with any given particle converges to 0 with a . The result holds also with less restrictive assumptions on the initial distribution.

PROPOSITION 7.1. *Let $\int_{\mathbb{R}^d} f_0(\cdot, v) \, dv \in L^d(U)$. Then the expected number of overlapping particles at a given point u_0 converges to zero for $a \rightarrow 0$.*

Proof. The expected number of particles in ball B_a of radius a around a $u_0 \in U$ is given by $p(u_0) = n \int_{B_a(u_0)} \int_{\mathbb{R}^d} df_0(u, v)$, which by the scaling and (2.8) can be estimated using the Hölder inequality

$$\begin{aligned} p(u_0) &\leq n \left(\int_{B_a} 1 \, du \right)^{(d-1)/d} \left(\int_{B_a} \left(\int_{\mathbb{R}^d} dv f_0(u, v) \right)^d \, du \right)^{1/d} \\ &= a^{1-d} (\kappa_{d+1} a^d)^{(d-1)/d} \left(\int_{B_a} \left(\int_{\mathbb{R}^d} dv f_0(u, v) \right)^d \, du \right)^{1/d} \rightarrow 0 \end{aligned}$$

for $a \rightarrow 0$ as $\int_{\mathbb{R}^d} dv f_0(\cdot, v) \in L^d(U)$. □

Whereas for $\int_{\mathbb{R}^d} dv f_0(u, v) = |u|^p$ near $u = 0$ with $-d < p < -1$, we still have $\int_{\mathbb{R}^d} dv f_0(u, v) \in L^1(U)$ but the expected number of particles in a ball of radius a around 0 tends to infinity; this effect will not be statistically relevant, as the growth is sublinear in n .

We now modify this example such that the expected number of nodes in the empirical trees tends to infinity for a tending to zero, even for short times when $f_0 \in L^1_{1+|v|}(U \times \mathbb{R}^d)$. Note that the idealized theory leads to the integral equation (4.35), which is well-defined for initial data in $L^1_{1+|v|}(U \times \mathbb{R}^d)$ and can be easily interpreted for measures. While the first example does not show nonvalidity due to singularities, it highlights difficulties in a proof for more general initial distributions, as tightness (4.31) was crucial to restrict the error estimates to trees of finite size.

PROPOSITION 7.2. *There exists an initial distribution $f_0 \in L^1_{1+|v|}(U \times \mathbb{R}^d)$ such that the expected number of overlaps $\int_{U \times \mathbb{R}^d} p(u) f_0(u, v)$ is unbounded as $a \rightarrow 0$.*

Proof. Let $(u_i)_{i \in \mathbb{N}^d}$ be an ordering of all dyadic fractions on the torus U such that for every pair i, j with

$$u_i = \left(\frac{i_1}{2^{k_1}}, \dots, \frac{i_d}{2^{k_d}} \right) \text{ with } \gcd(i_1, 2^{k_1}) = \dots = \gcd(i_d, 2^{k_d}) = 1,$$

$$u_j = \left(\frac{j_1}{2^{l_1}}, \dots, \frac{j_d}{2^{l_d}} \right) \text{ with } \gcd(j_1, 2^{l_1}) = \dots = \gcd(j_d, 2^{l_d}) = 1,$$

and $\max\{k_1, \dots, k_d\} > \max\{l_1, \dots, l_d\}$ then $i > j$. We consider

$$(7.1) \quad f_0(u, v) = c \sum_{j=1}^{\infty} c_j |u - u_j|^p \bar{f}(v)$$

with $-d < p < -1$ to be chosen later. The density $\bar{f} \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ is non-negative and normalized ($\int_{\mathbb{R}^d} d\bar{f} = 1$). The constant c and the sequence $(c_j)_{j \in \mathbb{N}}$ are chosen such that $\int_{U \times \mathbb{R}^d} df_0 = 1$. We will consider in particular the cases $d \geq 3$. Proposition 4.9 yields existence of a solution to (4.35), which is at least $L^1_{1+|v|}(U \times \mathbb{R}^d)$, when $f_0 \in L^1_{1+|v|}(U \times \mathbb{R}^d)$. For f_0 as in (7.1) there exists a $p \in (-d, -1)$ such that the empirical expected number of overlapping particles is unbounded.

The expected number of particles overlapping the first particle can be expressed as $\int_{U \times \mathbb{R}^d} df_0(u, v) p(u)$. If $u \in B_{a/2}(u_j)$, then for some constant C independent of j and a ,

$$p(u) = n \int_{B_a(u) \times \mathbb{R}^d} df_0(u', v) \geq nc \int_{B_a(u)} c_j |u' - u_j|^p du'$$

$$\geq nC \int_{B_{a/2}(u_j)} c_j |u' - u_j|^p du' \geq a^{1-d} C c_j \int_0^{a/2} r^p r^{d-1} dr = C c_j a^{1+p}.$$

We choose $J(a) \in \mathbb{N}$ such that the balls $B_{a/2}(u_j)$ are disjoint for $j = 1, \dots, J$. The expected number of overlaps can be bounded from below by

$$\int_{U \times \mathbb{R}^d} df_0(u, v) p(u) \geq \sum_{j=1}^J \int_{B_{a/2}(u_j) \times \mathbb{R}^d} df_0(u, v) p(u)$$

$$\geq \sum_{j=1}^J \left(\frac{a}{2}\right)^d c_j \left(\frac{a}{2}\right)^p C c_j a^{1+p} = C a^{d+2p+1} \sum_{j=1}^J c_j^2,$$

where C is a constant that does not depend on a . For $-d < p < -\frac{d+1}{2}$ this is unbounded as $a \rightarrow 0$. \square

We give now an example of nonvalidity if we allow concentrations in space and velocity space. Note that in the spatially homogeneous case we could prove validity unless there was concentration on lines within velocity space [MT10]. With \mathcal{H}^2 denoting the two-dimensional Hausdorff measure, we consider initial data concentrated on a sphere $S = \{u \in U : |u - u_0| = r\}$, where $r \in (\frac{1}{4}, \frac{1}{2})$ and $u_0 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$,

$$(7.2) \quad f_0(u, v) = c\mathcal{H}^2|_S \times \delta_0(v - (u - u_0)),$$

where c is a normalization constant.

PROPOSITION 7.3. *Let f_0 be given by (7.2) and let the solution of the gainless Boltzmann equation (2.5) be given by (4.35); then the convergence in probability (2.9) does not hold.*

Proof. We first observe for the idealized prediction that $f_t(u, v) \leq f_0(u - tv, v)$ by (4.35). Now suppose $f_0(u - tv, v) \neq 0$ for some $t \in (0, 1)$; then $f_0(u - (t - s)v, v') = 0$ for all $v' \neq v$ and all $0 < s < t$. Then $f_s(\tilde{u}, v') = 0$ for all $\tilde{u} = u - (t - s)v$ and all $v' \neq v$ by (4.35). Hence the integral in the argument of the exponential in (4.35) vanishes and $f_t(u, v) = f_0(u - tv, v)$. This is equivalent to pure transport until $t = 1$. The measure is concentrated on a decreasing sphere in u and the two-dimensional Hausdorff measure is scaled by $(\frac{r(t)}{r(0)})^2$.

On the other hand the hard sphere flow is well-defined with f_0 being a general measure. The n particles are distributed randomly with uniform distribution on a shrinking two-dimensional sphere of radius $r(1 - t)$ and surface area $4\pi r^2(1 - t)^2$.

If an iid distribution is used, then a macroscopic portion of the particles is instantaneously removed almost surely thanks to the definition of the scattering state $\beta_i^{(a)}(0)$ in (2.4). The following calculation shows that even in the case of more general distributions a macroscopic proportion of the particles undergoes a collision before time $t = 1$.

The surface covered by n balls of diameter a is $na^2\pi/4 = \pi/4$ by (2.8), i.e., at most the fraction $\frac{4\pi r^2(1-t)^2}{\pi/4} = 16r^2(1-t)^2$ of particles has not collided by time $t \in (0, 1)$. Thus for any empirical distribution $\frac{1}{n}\#\{i \mid (u_i(t), v_i(t)) \in U \times \mathbb{R}^d, \beta_i^{(a)}(t) = 1\} \leq 1/4$ for $3/4 < t < 1$. \square

REFERENCES

- [Ag84] V. AGOSHKOV, *Spaces of functions with differential-difference characteristics and the smoothness of solutions of the transport equation*, Dokl. Akad. Nauk SSSR, 276 (1984), pp. 1289–1293.
- [Ba77] J. BALL, *Strongly continuous semigroups, weak solutions, and the variation of constants formula*, Proc. Amer. Math. Soc., 63 (1977), pp. 370–373.
- [BGL93] C. BARDOS, F. GOLSE, AND C. LEVERMORE, *Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation*, Comm. Pure Appl. Math., 46 (1993), pp. 667–753.
- [BBS83] C. BOLDRIGHINI, L.A. BUNIMOVICH, AND Y.G. SINAI, *On the Boltzmann equation for the Lorentz gas*, J. Statist. Phys., 32 (1983), pp. 477–501.
- [CIP94] C. CERCIGNANI, R. ILLNER, AND M. PULVIRENTI, *The Mathematical Theory of Dilute Gases*, Appl. Math. Sci. 106, Springer-Verlag, New York, 1994.
- [CDPTW03] F. COPPEX, M. DROZ, J. PIASECKI, E. TRIZAC, AND P. WITTEWER, *Some exact results for Boltzmann's annihilation dynamics*, Phys. Rev. E, 79 (2003), p. 21103.
- [DP91] E. DAWSON AND E. PERKINS, *Historical Processes*, American Mathematical Society, Providence, RI, 1991.
- [DM78] C. DELLACHERIE AND P. MEYER, *Probabilities and Potential*, North-Holland Math. Stud. 29. North-Holland, Amsterdam, 1978.

- [DP01] R. DeVORE AND G. PETROVA, *The averaging lemma*, J. Amer. Math. Soc., 14 (2001), pp. 279–296.
- [DL89a] R. DiPERNA AND P.L. LIONS, *On the Cauchy problem for Boltzmann equations: Global existence and weak stability*, Ann. Math., 130 (1989), pp. 321–366.
- [DL89b] R. DiPERNA AND P.L. LIONS, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math., 98 (1989), pp. 511–547.
- [DLM91] R. DiPERNA, P.L. LIONS, AND Y. MEYER, *L^p regularity of velocity averages*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 8 (1991), pp. 271–287.
- [DFPR95] M. DROZ, L. FRACHEBOURG, J. PIASECKI, AND P.-A. REY, *Ballistic annihilation kinetics for a multiveLOCITY one-dimensional ideal gas*, Phys. Rev. E, 51 (1995), pp. 5541–5548.
- [EF85] Y. ELSKENS AND H. FRISCH, *Annihilation kinetics in the one-dimensional ideal gas*, Phys. Rev. A, 31 (1985), pp. 3812–3816.
- [Ge90] P. GÉRARD, *Moyennisation et régularité deux-microlocale*, Ann. Sci. École Norm. Sup., 23 (1990), pp. 89–121.
- [GG92] P. GÉRARD AND F. GOLSE, *Averaging regularity results for PDEs under transversality assumptions*, Comm. Pure Appl. Math., 45 (1992), pp. 1–26.
- [GLPS99] F. GOLSE, P.-L. LIONS, B. PERTHAME, AND R. SENTIS, *Regularity of the moments of the solution of a transport equation*, J. Funct. Anal., 76 (1988), pp. 110–125.
- [GS04] F. GOLSE AND L. SAINT-RAYMOND, *The Navier-Stokes limit of Boltzmann equation: Convergence proof*, Invent. Math., 155 (2004), pp. 81–161.
- [HR07] A. HAMMOND AND F. REZAKHANLOU, *The Kinetic Limit of a System of Coagulating Brownian Particles*, Arch. Ration. Mech. Anal., 185 (2007), pp. 1–67.
- [IP89] R. ILLNER AND 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), pp. 143–146.
- [KS88] J. KRUG AND H. SPOHN, *Universality classes for deterministic surface growth*, Phys. Rev. A, 38, (1988), pp. 4271–4283.
- [Lan75] O. LANFORD, *Time Evolution of Large Classical Systems*, J. Moser, ed., Lecture Notes in Phys. 38, Springer-Verlag, New York, 1975, pp. 1–113.
- [LN80] R. LANG AND X. NGUYEN, *Smoluchowski’s theory of coagulation holds rigorously in the Boltzmann-Grad limit*, Z. Wahrs. Verw. Geb., 54 (1980), pp. 227–280.
- [LM01a] P.-L. LIONS AND N. MASMOUDI, *From the Boltzmann equations to the equations of incompressible fluid mechanics. I*, Arch. Ration. Mech. Anal., 158 (2001), pp. 173–193.
- [LM01b] P.-L. LIONS AND N. MASMOUDI, *From the Boltzmann equations to the equations of incompressible fluid mechanics. II*, Arch. Ration. Mech. Anal., 158 (2001), pp. 195–211.
- [MT08] K. MATTHIES AND F. THEIL, *Validity and non-validity of propagation of chaos*, in Analysis and Stochastics of Growth Processes and Interface Models, P. Mörters, R. Moser, M. Penrose, H. Schwetlick, and J. Zimmer, eds., Oxford University Press, Oxford, UK, 2008, pp. 101–119.
- [MT10] K. MATTHIES AND F. THEIL, *Validity and failure of the Boltzmann approximation of kinetic annihilation*, J. Nonlinear Sci., 20 (2010), pp. 1–46.
- [Per04] B. PERTHAME, *Mathematical tools for kinetic equations*, Bull. Amer. Math. Soc. (N.S.), 41 (2004), pp. 205–244.
- [Paz83] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Appl. Math. Sci., Springer-Verlag, New York, 1983.
- [Pia95] J. PIASECKI, *Ballistic annihilation in a one-dimensional fluid*, Phys. Rev. E, 51 (1995), pp. 5535–5540.
- [PTD02] J. PIASECKI, E. TRIZAC, AND M. DROZ, *Dynamics of ballistic annihilation*, Phys. Rev. E, 65 (2002), p. 66111.
- [Rez04] F. REZAKHANLOU, *Boltzmann-Grad limits for stochastic hard sphere models*, Comm. Math. Phys., 248 (2004), pp. 553–637.
- [Szn87] A. SZNITMAN, *Propagation of chaos for a system of annihilating Brownian spheres*, Comm. Pure Appl. Math., 40 (1987), pp. 663–690.
- [Spo78] H. SPOHN, *The Lorentz process converges to a random flight process*, Comm. Math. Phys., 60 (1978), pp. 277–290.
- [Spo91] H. SPOHN, *Large Scale Dynamics of Interacting Particles*, Texts Monogr. Phys., Springer-Verlag, New York, 1991.